



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 180127

TO: Phuong Bui
Location: REM-2a!5&2C15 *2C18*
Art Unit: 1638
Tuesday, August 24, 2004

Case Serial Number: 09/900237

From: David Schreiber
Location: Biotech-Chem Library
Remsen E01A61
Phone: 272-2526

david.schreiber@uspto.gov

Search Notes

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2004, 01:11:20 ; Search time 13694 Seconds
(without alignments)
11476.694 Million cell updates/sec

Title: US-09-900-237A-29
Perfect score: 3626
Sequence: 1 gcaagagaaccccgctcca.....tcacatttggaggagtttt 3626

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3626	100.0	3626	8	BT009438	BT009438 Triticum
2	2694.4	74.3	3954	8	AK072356	AK072356 Oryza sat
3	2637.8	72.7	3264	6	AX653232	AX653232 Sequence
4	2595	71.6	3795	8	AF200533	AF200533 Zea mays
5	2595	71.6	3799	6	AX338680	AX338680 Sequence
6	2572.2	70.9	3745	8	AF200528	AF200528 Zea mays
7	2360.8	65.1	4282	8	AK069196	AK069196 Oryza sat
8	2352.4	64.9	3222	6	AX652952	AX652952 Sequence
9	2325.2	64.1	3676	8	AF200529	AF200529 Zea mays
10	1720	47.4	3532	8	AY055724	AY055724 Populus t
11	1699.8	46.9	3723	8	AF150630	AF150630 Gossypium
12	1679.4	46.3	3229	8	BT002335	BT002335 Arabidops
13	1675.2	46.2	3682	8	AF027174	AF027174 Arabidops
14	1673.6	46.2	3614	6	AR267559	AR267559 Sequence
15	1673.6	46.2	3614	6	AX030946	AX030946 Sequence
16	1673.6	46.2	3614	6	BD022678	BD022678 Manipulat
17	1412.4	39.0	3732	8	AK099228	AK099228 Oryza sat
18	1412.4	39.0	3768	8	AK102140	AK102140 Oryza sat
19	1412.4	39.0	3801	8	AK099281	AK099281 Oryza sat
20	1412.4	39.0	3802	8	AK067967	AK067967 Oryza sat
21	1410.8	38.9	3897	8	AK100188	AK100188 Oryza sat
22	1409.8	38.9	3764	8	AK098978	AK098978 Oryza sat
23	1408.4	38.8	3725	8	AF200526	AF200526 Zea mays
24	1383.6	38.2	3851	6	BD236020	BD236020 Materials
25	1373	37.9	3752	8	AF200525	AF200525 Zea mays
26	1369.4	37.8	3640	8	AY162181	AY162181 Populus t
27	1343.4	37.0	3603	6	AX030942	AX030942 Sequence
28	1343.4	37.0	3603	6	BD022676	BD022676 Manipulat
29	1343.4	37.0	3603	8	BT008654	BT008654 Arabidops
30	1343.4	37.0	3763	8	BT008654	BT008654 Arabidops
31	1341.8	37.0	3673	6	AR267560	AR267560 Sequence
32	1341.8	37.0	3673	6	AX030948	AX030948 Sequence
33	1341.8	37.0	3673	6	BD022679	BD022679 Manipulat
34	1340.2	37.0	3812	8	AF200532	AF200532 Zea mays
35	1331.4	36.7	2828	8	AF200527	AF200527 Zea mays
36	1331.4	36.7	2830	6	AX338676	AX338676 Sequence
37	1316.8	36.3	3538	8	AF200530	AF200530 Zea mays
38	1314.6	36.3	3746	8	AK100914	AK100914 Oryza sat
39	1294.6	35.7	4127	8	AK121193	AK121193 Oryza sat
40	1281.2	35.3	4029	8	AK100877	AK100877 Oryza sat
41	1280.4	35.3	4208	8	AK073561	AK073561 Oryza sat
42	1271	35.1	3968	8	AF200531	AF200531 Oryza sat
43	1246.2	34.4	3631	8	AK121170	AK121170 Oryza sat
44	1244.8	34.3	3847	8	AF304374	AF304374 Nicotiana
45	1235.6	34.1	3221	8	AK067386	AK067386 Oryza sat

ALIGNMENTS

RESULT 1
BT009438
LOCUS
DEFINITION Triticum aestivum clone wlmk4.pk0015.all:fls, full insert mRNA
3626 bp mRNA linear PLN 20-JUN-2003

ACCESSION BT009438
VERSION BT009438.1 GI:32128989
KEYWORDS FLI CDNA.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 3626)
AUTHORS Tingey,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H.,

Caraher, N.R., Hanafey, M.K. and Hainey, C.F.

TITLE Direct Submission
JOURNAL Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA

FEATURES Location/Qualifiers

source
1..3626
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wlmk4.pk0015.all.fis"

ORIGIN

Query Match 100.0%; Score 3626; DB 8; Length 3626;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACAGAGAACCCGCTCCAGCTCTCTCGTGGTGGGTTGGATCGCTCTGCGCGCCA 60
Db 1 GCACAGAGAACCCGCTCCAGCTCTCTCGTGGTGGGTTGGATCGCTCTGCGCGCCA 60

QY 61 TGGACGGCGACGGGACGGCTTGAAGTCCGGGAGGACGGGGCCGGGAGCTGTGCCAGA 120
Db 61 TGGACGGCGACGGGACGGCTTGAAGTCCGGGAGGACGGGGCCGGGAGCTGTGCCAGA 120

QY 121 TCTGCGCGACGGCTGGGACACACCTTGGACGGGACGCTTTCACGGCTCGGAGTCT 180
Db 121 TCTGCGCGACGGCTGGGACACACCTTGGACGGGACGCTTTCACGGCTCGGAGTCT 180

QY 181 GCGCTTCCCGCTCTGCGCGCTCTGTACGAGCAGGACGAGGAGGACCCAGGCGCT 240
Db 181 GCGCTTCCCGCTCTGCGCGCTCTGTACGAGCAGGACGAGGAGGACCCAGGCGCT 240

QY 241 GCCTCCAGTCAAGACCAAGTACAAGCCACAGAGGAGCCAGCGATCCGCGGGAGG 300
Db 241 GCCTCCAGTCAAGACCAAGTACAAGCCACAGAGGAGCCAGCGATCCGCGGGAGG 300

QY 301 AAGCGACGACCTGATGCGCGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 301 AAGCGACGACCTGATGCGCGATGATGATGATGATGATGATGATGATGATGATGAT 360

QY 361 AGGACAGAGCAGAGATGCTGACAGGATGCGGAGCTGGCGATGAACACCGGGGCA 420
Db 361 AGGACAGAGCAGAGATGCTGACAGGATGCGGAGCTGGCGATGAACACCGGGGCA 420

QY 421 GTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGGAGATCGGCTCTCCAAATGACA 480
Db 421 GTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGGAGATCGGCTCTCCAAATGACA 480

QY 481 GTGGAGATCCCTAGGGGATAGTCCCTTCAGTCAACACAGCCAGATGTCAGGAGAA 540
Db 481 GTGGAGATCCCTAGGGGATAGTCCCTTCAGTCAACACAGCCAGATGTCAGGAGAA 540

QY 541 TCCTCGAGCTTCGCTGATCATCATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 541 TCCTCGAGCTTCGCTGATCATCATGATGATGATGATGATGATGATGATGATGATGAT 600

QY 601 CTGCTTCCCTATGTAATCATTCAACAAATCGGTCAAGGAGTTCCTCGGAGATTTG 660
Db 601 CTGCTTCCCTATGTAATCATTCAACAAATCGGTCAAGGAGTTCCTCGGAGATTTG 660

QY 661 GGAATGTTGCTGGAAGAGAGAGTTGATGGCTGGAATAATGAAGCAGACAGGGTCCGA 720
Db 661 GGAATGTTGCTGGAAGAGAGAGTTGATGGCTGGAATAATGAAGCAGACAGGGTCCGA 720

QY 721 TTCCATGACTTAATGGACAAAGCAATTCCTCTGTAAGGTTCGGGACAGTACTGACATCG 780
Db 721 TTCCATGACTTAATGGACAAAGCAATTCCTCTGTAAGGTTCGGGACAGTACTGACATCG 780

QY 781 ATGCATCTACTGAATACAAATGGAAGCGCTTTACTGAATGATGAATGATGAATGATGA 840
Db 781 ATGCATCTACTGAATACAAATGGAAGCGCTTTACTGAATGATGAATGATGAATGATGA 840

QY 841 TATCTAGAAAAGTCCCATTTGCTTCTCCAAAATAAATCCCTACAGAAATGGTCAATGTTTC 900
Db 841 TATCTAGAAAAGTCCCATTTGCTTCTCCAAAATAAATCCCTACAGAAATGGTCAATGTTTC 900

QY 901 TCGCGTTGGTTGTTCTAAGCATCTTCTGTCGATACCGTCTCAAAATCTCTGCGTAAATG 960
Db 901 TCGCGTTGGTTGTTCTAAGCATCTTCTGTCGATACCGTCTCAAAATCTCTGCGTAAATG 960

QY 961 CATACCCAGTGGCTTTTATCTGTATATGTAGATTTGGTTTGGTTTATCTCTGATAC 1020
Db 961 CATACCCAGTGGCTTTTATCTGTATATGTAGATTTGGTTTGGTTTATCTCTGATAC 1020

QY 1021 TGGATCAGTTCCCGAAGTGGTTTCCAATCAACCGGAGACCTACCTTGTATAGACTGGCTT 1080
Db 1021 TGGATCAGTTCCCGAAGTGGTTTCCAATCAACCGGAGACCTACCTTGTATAGACTGGCTT 1080

QY 1081 TAAGGTATCACCGAGAGGTGAACCGTCTCAGTTGGCTGCTGTGTGACATATTTGTGAGTA 1140
Db 1081 TAAGGTATCACCGAGAGGTGAACCGTCTCAGTTGGCTGCTGTGTGACATATTTGTGAGTA 1140

QY 1141 CAGTCCAGCCCTTGAAGGAGGCCACCTATCTGCTCACTGCCAACACTGTGCTATCCATTCTTG 1200
Db 1141 CAGTCCAGCCCTTGAAGGAGGCCACCTATCTGCTCACTGCCAACACTGTGCTATCCATTCTTG 1200

QY 1201 CTGTTGATTTATCCCGTGGACAAAGTCTCTTGTCTATCTATCTGATGACGAGCTTCAATGC 1260
Db 1201 CTGTTGATTTATCCCGTGGACAAAGTCTCTTGTCTATCTATCTGATGACGAGCTTCAATGC 1260

QY 1261 TGAATTTTGAAGCATTTGGCTGAGACTTCTGAGTTTCTAGGAAATGGGTACCAATTTGTGA 1320
Db 1261 TGAATTTTGAAGCATTTGGCTGAGACTTCTGAGTTTCTAGGAAATGGGTACCAATTTGTGA 1320

QY 1321 AGAAGTATGACATTGAACCCAGAGCTCCGAGTTTCTTCTTTCAGAGAAATTTGATTACC 1380
Db 1321 AGAAGTATGACATTGAACCCAGAGCTCCGAGTTTCTTCTTTCAGAGAAATTTGATTACC 1380

QY 1381 TGAAGACAAAGTCCAGCTTCAATTTGTTAAAGACCGCGGGCCATGAAGAGAGATATG 1440
Db 1381 TGAAGACAAAGTCCAGCTTCAATTTGTTAAAGACCGCGGGCCATGAAGAGAGATATG 1440

QY 1441 AGAATTTAAATCAGATTAATGCGCTAGTTTCTTAAGCATTTGAAGTCCCGAGGAG 1500
Db 1441 AGAATTTAAATCAGATTAATGCGCTAGTTTCTTAAGCATTTGAAGTCCCGAGGAG 1500

QY 1501 GATGGATCATGCAAGATGGCACCATGGCCAGGAAACAAATACAGGGATCATCTCGGAA 1560
Db 1501 GATGGATCATGCAAGATGGCACCATGGCCAGGAAACAAATACAGGGATCATCTCGGAA 1560

QY 1561 TGATTCAGGTTTCTTGGTCAAGTGGTGGCTTGTATCTGATGAGGGTATGAGCTCCGCC 1620
Db 1561 TGATTCAGGTTTCTTGGTCAAGTGGTGGCTTGTATCTGATGAGGGTATGAGCTCCGCC 1620

QY 1621 GTTTAGTTATGTTCTCGTGAAGAGGCTCTGGTTCCAGCACCACAGAGGCTGGTG 1680
Db 1621 GTTTAGTTATGTTCTCGTGAAGAGGCTCTGGTTCCAGCACCACAGAGGCTGGTG 1680

QY 1681 CCATGAATGCCCTTGTTCGTGTCTCAGTCTCTTACTTAATGGACATAATGTTGATTC 1740
Db 1681 CCATGAATGCCCTTGTTCGTGTCTCAGTCTCTTACTTAATGGACATAATGTTGATTC 1740

QY 1741 TTGATTTGATCATCATCAACAGCAGGAGCTGTCCGAGAGCTATGTGCTTCTCTAA 1800
Db 1741 TTGATTTGATCATCATCAACAGCAGGAGCTGTCCGAGAGCTATGTGCTTCTCTAA 1800

QY 1801 TGGATCAAACTAGGTCCGCAAGTCTGTTATGTGAGTTTCCCAAAAGGTTTGTATGGGA 1860
Db 1801 TGGATCAAACTAGGTCCGCAAGTCTGTTATGTGAGTTTCCCAAAAGGTTTGTATGGGA 1860

QY 1861 TTGATAGGAATGATTCGATATGCAAAAGGAACTGTCTTTTGTATTAATTTAGTTGAGG 1920
Db 1861 TTGATAGGAATGATTCGATATGCAAAAGGAACTGTCTTTTGTATTAATTTAGTTGAGG 1920

QY 1921 GCCTTGACGGCATTCAGAGGACCGATTATGTTGGAACTGGTTGTTTTCACAGAACAG 1980

Db 1921 GCTTGAAGGCAATCAAGGACAGTTTATGTTGGAACTGTTGTTTCAACAGACAG 1980
QY 1981 CTATCTAGTTTATGAGCCGCCAAATTAAGCGGAAGACAGGTTTCTTGGGATCACTAT 2040
Db 1981 CTATCTAGTTTATGAGCCGCCAAATTAAGCGGAAGACAGGTTTCTTGGGATCACTAT 2040
QY 2041 GTGGGGCAAGAAAGGCAAGCAAGTCAAGAAAGAGGAGCTCAGATAGAAAGTCA 2100
Db 2041 GTGGGGCAAGAAAGGCAAGCAAGTCAAGAAAGAGGAGCTCAGATAGAAAGTCA 2100
QY 2101 ACAAGCATGTGACAGTTCTGTTCAGTATCAATCTCGAAGACATAGAGGAGGTTG 2160
Db 2101 ACAAGCATGTGACAGTTCTGTTCAGTATCAATCTCGAAGACATAGAGGAGGTTG 2160
QY 2161 AAGGTGCTGGTTTATGATGAGAAATCAGTTCTCATGTCTCAATGCTCAGATAGAGAA 2220
Db 2161 AAGGTGCTGGTTTATGATGAGAAATCAGTTCTCATGTCTCAATGCTCAGATAGAGAA 2220
QY 2221 GATTTGGCCAGTCAGCAGCAATTTGCTCCACTCTCATGGAATATGTTGCTTCCTC 2280
Db 2221 GATTTGGCCAGTCAGCAGCAATTTGCTCCACTCTCATGGAATATGTTGCTTCCTC 2280
QY 2281 AGTCTCCACTCCAGATCTCTTTGAAAGAGCTATCCATGTCTAATAGTTGTCGCTATG 2340
Db 2281 AGTCTCCACTCCAGATCTCTTTGAAAGAGCTATCCATGTCTAATAGTTGTCGCTATG 2340
QY 2341 AGGCAAGTCTGAATGGGGAATGAGTTGGTGGATCTATGGATCTGTGTCAGAGATA 2400
Db 2341 AGGCAAGTCTGAATGGGGAATGAGTTGGTGGATCTATGGATCTGTGTCAGAGATA 2400
QY 2401 TTTCTACTGGATTCAAGATGACGCAAGAGGCTGCGCTTCACTCTATTGCAATGCCAAGC 2460
Db 2401 TTTCTACTGGATTCAAGATGACGCAAGAGGCTGCGCTTCACTCTATTGCAATGCCAAGC 2460
QY 2461 GCCAGCTTTCAAGGATCTGCCCCCATCAATCTTTAGATCGTCTGCAACCAAGTCTGC 2520
Db 2461 GCCAGCTTTCAAGGATCTGCCCCCATCAATCTTTAGATCGTCTGCAACCAAGTCTGC 2520
QY 2521 GGTGGGCTCTGGTCTCTGTAATCTTTTTCAGCGGAGATGCGCCCTATGTTAGTGGT 2580
Db 2521 GGTGGGCTCTGGTCTCTGTAATCTTTTTCAGCGGAGATGCGCCCTATGTTAGTGGT 2580
QY 2581 ACGAGGGGCGCTCAAGTTCTGGAGATGCGCTTACATCAACACACCACTTTACCCAC 2640
Db 2581 ACGAGGGGCGCTCAAGTTCTGGAGATGCGCTTACATCAACACACCACTTTACCCAC 2640
QY 2641 TAACTCTCTCCGCTCTAGTCTATTGATGATGATGATGATGATGATGATGATGATGAT 2700
Db 2641 TAACTCTCTCCGCTCTAGTCTATTGATGATGATGATGATGATGATGATGATGATGAT 2700
QY 2701 AGTTCAATGCGAGATGAGCACTTGGGCAATCTGGGCAATCTGGGCAATCTGGGCAAT 2760
Db 2701 AGTTCAATGCGAGATGAGCACTTGGGCAATCTGGGCAATCTGGGCAATCTGGGCAAT 2760
QY 2761 CAATTTTCCGCTCTAGTCTATTGATGATGATGATGATGATGATGATGATGATGAT 2820
Db 2761 CAATTTTCCGCTCTAGTCTATTGATGATGATGATGATGATGATGATGATGATGAT 2820
QY 2821 GGAGGAATGAACAGTTCTGGGCAATCTGGGCAATCTGGGCAATCTGGGCAATCTGGG 2880
Db 2821 GGAGGAATGAACAGTTCTGGGCAATCTGGGCAATCTGGGCAATCTGGGCAATCTGGG 2880
QY 2881 AGGTCTCTCAAGTGTCTGCGGATGCGACCACTTCACTGTCTACCTCAAGGCTA 2940
Db 2881 AGGTCTCTCAAGTGTCTGCGGATGCGACCACTTCACTGTCTACCTCAAGGCTA 2940
QY 2941 ATGAGGAGAGAGGACCTTCTGAGCTCTACATCTTCAAGTGGACGAGCTTCTCATCC 3000
Db 2941 ATGAGGAGAGAGGACCTTCTGAGCTCTACATCTTCAAGTGGACGAGCTTCTCATCC 3000
QY 3001 CTCCGACACCAATTTGATCATTAACATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3060

Db 3001 CTCCGACACCAATTTGATCATTAACATGTTGGTGGTGGTGGTGGTGGTGGTGGTGG 3060
QY 3061 TCAACAGTGGTTACCAATCAGGGGCGCTCTTTGGGAAGCTCTTTTGGCTTTCTGGG 3120
Db 3061 TCAACAGTGGTTACCAATCAGGGGCGCTCTTTGGGAAGCTCTTTTGGCTTTCTGGG 3120
QY 3121 TGATTTGTTCACTATATACCAATCTCTCAAGGTCCTATGGGAGGAAACCGCACCGA 3180
Db 3121 TGATTTGTTCACTATATACCAATCTCTCAAGGTCCTATGGGAGGAAACCGCACCGA 3180
QY 3181 CGATTTGTCATCTGCTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3240
Db 3181 CGATTTGTCATCTGCTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3240
QY 3241 TTGATCCATTTCACTACCCGCTCTGCTGGGCCCAATATCCAAACCTGTGGCATCACTGCT 3300
Db 3241 TTGATCCATTTCACTACCCGCTCTGCTGGGCCCAATATCCAAACCTGTGGCATCACTGCT 3300
QY 3301 AGGAAGTGGGAGTTTGTAGAGACAGAAATATACAGTATCGAGCAACACCGCGGA 3360
Db 3301 AGGAAGTGGGAGTTTGTAGAGACAGAAATATACAGTATCGAGCAACACCGCGGA 3360
QY 3361 GCCAGAGATATTTATGTTGGGTTGTCAATTAATCTGAGAAAGTTGTCAAAATG 3420
Db 3361 GCCAGAGATATTTATGTTGGGTTGTCAATTAATCTGAGAAAGTTGTCAAAATG 3420
QY 3421 AGAAACACATTTGTAATAGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
Db 3421 AGAAACACATTTGTAATAGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
QY 3481 TTTCTTTTTCGAAACAAAGGATCTCATTTGTAACCTATAGGAATTTTCTATGAGGA 3540
Db 3481 TTTCTTTTTCGAAACAAAGGATCTCATTTGTAACCTATAGGAATTTTCTATGAGGA 3540
QY 3541 CTTTGGATTTGAGGAATGCACTATGAAATGTTGATTTATTTATTTATTTATTTATTT 3600
Db 3541 CTTTGGATTTGAGGAATGCACTATGAAATGTTGATTTATTTATTTATTTATTTATTT 3600
QY 3601 TGTCTTCCATTTTGGAGGAGTTT 3626
Db 3601 TGTCTTCCATTTTGGAGGAGTTT 3626

RESULT 2

AK072356

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) cDNA clone J023059I02, full

insert sequence.

ACCESSION

AK072356

VERSION

FLI CDNA; CAP trapper.

KEYWORDS

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1

REFERENCE

AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team.,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Ohtsuki, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imorani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,

1681 CTTGATTGATCACTACATCAACACAGAGGCTCTCGGGAAGCTATGCTTCCTT 1740
1800 ATGGATCCAAACCTAGGTCGGAGTCTGTTATGTCAGTTCCTCCACAAAGGTTGATGG 1859
1741 ATGGATCCAAACCTAGGAGGAGTGTCTGTTATGTTCAAGTTCCTCCACAAAGGTTGATGGT 1800
1860 ATTGATAGGAATGATCATATGCAACAGCAACACACTGTCTTTTTCATATTAACCTTGAGG 1919
1801 ATTGATAGGAATGATCATATGGAACAGCAACACTGTCTTTTTCATATTAACCTTGAGG 1860
1920 GGCCTTGACGGCAATCAAGGACAGTTTATGTTGGAACCTGGTGTGTTTCAACAGAA 1979
1861 GGTCTTGATGGCATCCAGGACAGTTTATGTTGGAACCTGGTGTGTTTCAACAGAACT 1920
1980 GCTATCTATGTTATGAGCCCCCAATTA ---AGGCGAAGAACGAGTTCCTTGCAATCA 2036
1921 GCTCTATATGTTATGAAACCCCAATTAAGCAGAAAGAAAGGAAAGTTCTTGTGTCATCA 1980
2037 CTATGTGGGGCAAGAAGAGGCAAGCAAGTCAAGAAAGAGGAGCTCAGATGAAGAAAG 2096
1981 CTATGTGGGGCAGAAAGAGGCAAGCAAGTCAAGAAAGAGGTTCCGACAAAGAAAG 2040
2097 TCGAACAGAGTGTGACAGAGTCTGTTCCAGTATTAATCTCGAAGACATAGAGAGGCT 2156
2041 TCAAAACAGACAGTGGACAGTCTGTCAGGTTTCAATCTTGAAGATATAGAGAGGCT 2100
2157 GTTGAAGTGTGGGTTTATGATGAGAAATCAGTCTCATGTCCTCAATGAGCTTAGAG 2216
2101 GTTGAAGTGTGGGTTTATGATGAGAAATCAGTCTTATGTCCTCAATGAGCTTAGAG 2160
2217 AAGAGATTGGGCAGTCAGCAGCATTTGTTGCTCCACTCTCATGGAATATGCTGTGTT 2276
2161 AAGAGATTGGGCAGTCAGCAGCTTTGTCAGGTTTGTGCTCCACTCTCATGGAATATGCTGTGTT 2220
2277 CTTAGTCTCTCACTCCAGAAATCTCTTTTGAAGAAGCTATCCATGTCATAAGTTGTGGC 2336
2221 CTTCAATCTGCAACCCCAAGAACTCTCTTTTGAAGAAGCTATCCATGTCATAAGTTGTGGC 2280
2337 TATGGGACAGTCTGATGGGAACTGA -----GATTGGTGGATC 2378
2281 TATGAGGCAAGACCGAATGGGAACTGAGGCTATGCAACTTGCATGTTGGTGGATC 2340
2379 TATGATCTGTACAGAAAGATATCTTACTGATTCAGATGACGACGAGGCTGGGCT 2438
2341 TACGTTCCGTCACAGAGATATCTCTAGTATTCAGATGATGATGCGGAGCTGGAGA 2400
2439 TCAGTCTATGATGCCCAAGCGCCAGCTTTCAAGGAGTCTGCCCCCATCAATCTTTCA 2498
2401 TCAATCTACTGATGCCCAAGCGCCAGCTTTCAAGGAGTCTGCCCTATCAATCTTTCA 2460
2499 GATCGTCTGAACCAAGTCTGCGGTGGGCTCTCGGTTCTGTTGAAATCTTTTTCAGCGG 2558
2461 GATCGTCTTAACCAAGTCTGCGGTGGGCACTGTTCTGTTGAAATCTTTTTCAGTCCG 2520
2559 CATTCGCCCTTATGATGGCTTACGAGGCGGCTCAAGTCTCTCGAGAGATTCGCTTAC 2618
2521 CATTTGCCATATGTAAGGCTATGAGGACGCTTAAAGTCTTTCGAGAGATTCGCTTAC 2580
2619 ATCAACACACATTTACCACTACCTCTCTCCGCTTCTAGTCTATTTGATATGCTT 2678
2581 ATCAACACACATTTATTCATTCATGATCGATCCGCTCTCATATATCTGTTGTTGCT 2640
2679 GCTATCTCTGCTCACTGGAAAGTTCATATGCGAGAGATAGCACTTGGCCAGATTC 2738
2641 GCTATCTGTTGCTCACTGGGAAGTTCATATCCAGAGATTAGCACTTGTCTAGTAT 2700
2739 TGGTTCATTGCGCTCTCTCTTTCAATTTTGGCACTGTGTTGCTTTCAGATGAGTGGAGT 2798
2701 TGGTTCATCTCTCTCTTCAATTTTGGCACTGTGTTGCTTTCAGATGAGTGGAGT 2760
2799 GGTGTGTCATTCAGAGTGTGGAGGAATGAACAGTTCTGCGTCAATGGAGGATCTCT 2858

2761 GGTGTGTCATCGATCAGTGTGGAGGATGAACAGTTCTGGGTATTGGAGGTATATCT 2820
2859 GCACATCTGTTGGCGCTCTTTCAAGGTCTTTGAAAGGTCTTCCGGTATCGACACCAAC 2918
2821 GGCATCTTTTGGCGCTCTTCCAGGCTCTCTCAAGGTCTTGGTATCGACACCAAT 2880
2919 TTCACTGTCACTCAAGGCTAATGACGAAGAGGAGCTTTTGGCTGAGCTCTACATGTTT 2978
2881 TTCACTGTCACTCAAGGCTTCTGATGAAGATGGGACTTTTGGCTGAGCTCTACATGTTT 2940
2979 AAGTGGACGACGCTTCTCATCCCTCCGACGACCAATTTTTCATTAACATGTTGTTGTC 3038
2941 AAGTGGACGACGCTTCTCATCCACGAGGACCACTTTGATCATTAACCTGTTGTTGTT 3000
3039 GTTGTGGGACCTCTCAAGGCTAATGACGAAGAGGAGCTTTTGGCTGAGCTCTACATGTTT 3098
3001 GTTGTGGGATCTCATACGCGATCAACGCGCTACCATTCATGGGAGCGCTCTTTGGC 3060
3099 AAGCTCTTCTTGGCTCTTGGGCTGATTTTCACTTATACCAATTCCTCAAGGCTCTTATG 3158
3061 AAGCTCTTCTTGGCTCTTGGGCTGATTTTCACTTATACCAATTCCTCAAGGCTCTTATG 3120
3159 GGCAGGCAAAACCGCACACCGACGATTTGTCATGCTGCGGCTGCTCTCGCTTCTATC 3218
3121 GGTGGGCAAAACCGCACCTCCGACCATCGTTGTTGTTGGCAATCTTCTGCTTCGATC 3180
3219 TTCTCTTCTGTTGGGTTGCTGTTGATCATTCATACCGGCTCTGCTGCGCCAAATATC 3278
3181 TTCTCATTTGCTGTTGGTTTGGCATCGATCGATTCACACCGGTTGTCACCGGCGCAGATACC 3240
3279 CAAACCTGTGGCATCAACTGCTA 3301
3241 CAAACATGGCATCAACTGCTA 3263

AF200533 3795 bp mRNA linear PIN 31-AUG-2000
Zea mays cellulose synthase-9 (Cesa-9) mRNA, complete cds.
AF200533
AF200533.1 GI:9622889

Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3795)
Holland, N., Holland, D., Helentjaris, T., Dhugga, K.S.,
Xocoostle-Cazares, B. and Delmer, D.P.,
A comparative analysis of the plant cellulose synthase (Cesa) gene
family
Plant Physiol. 123 (4), 1313-1324 (2000)
20398328
10938350
2 (bases 1 to 3795)
Dhugga, K.S. and Helentjaris, T.G.
Direct Submission
Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred
International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
Location/Qualifiers
1..3795
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/note="identified and isolated from the maize genome
database CDS collection of Pioneer Hi-Bred International,
Inc"
1..3795
/gene="Cesa-9"
238..3477
/gene="Cesa-9"
/note="Cesa-9"

gene
CDS

KRPAPKSGAPINLSRLNOVLWALGSEVILFSLRHCPWLVGYGRLKFLERAYINTT
IYPLTSLPLLYICILPAICLLKTKFIIPIEISNPASIFWISLISIPATGILEMRNSGV
GDEWRNRFQWIVIGISAHLFVQGLIKVLAGIDTNEFTVSKASDEGDDFAELIYMF
KWTLLIPTPTLLIINLVGVAGISYAINSGYQSMGLFGLFPAFWIVLIPYELKLG
LMGRQNRTEPTIVVWAILLASIFSLWLRIDPFTNRVTGPDTRTCGICN"

ORIGIN

Query Match 71.6%; Score 2595; DB 6; Length 3799;

Best Local Similarity 87.1%; Pred. No. 0;

Matches 2871; Conservative 0; Mismatches 420; Indels 4; Gaps 2;

QY	36	CGGGTGGATCGCTCGCGCGCATGGAGCGGACGCGACGCCCTCGAAGTCGGGAGG	95
DB	214	CGGTTCTGGCGAGCTCGCCTCGCATGGAGGGGACGCGACGGTGAAGTCGGGAGG	273
QY	96	CACGGGCGGGGACGCTGTGCCAGATCTGCCGACGGCCTGGGCCACCGATTGGACGGC	155
DB	274	CGCGGGGAGGGCAGGTGTGCCAGATCTGCCGCGATGGCGTGGGCATACGGCGGAGGGA	333
QY	156	GACGTCTTACCGCCTCGGACGCTCGCGCTTCCCGTCTGCCGCGCCCTGCTACGAGCAC	215
DB	334	GACGTCTTACCGCCTCGGACGCTCGCGGTTCCCGGTGTGCCGCGCCCTGCTACGAGTAC	393
QY	216	GAGCGAAGAGGGACCCGAGGCTGCTCCAGTGAAGACCAAGTACAAGCGCCACAGA	275
DB	394	GAGCGAAGAGGGACCAAGGTGCCCCAGTGCACAAACAAAGTACAAGCGCCACAGA	453
QY	276	GGGAGCCACGATCCCGGGGAGGAGGACGACACTGATCCGATGATGTAGTGAC	335
DB	454	GGGAGTCCAGCGATCCGAGGGGAGGAGGACGATCTGATCCGATGATGTAGCGAC	513
QY	336	TTCAACTACCTGATCTGACATCGGACGACGACGACGAGAGATGCTGACAGATGCGC	395
DB	514	TTCAACTACCTGATCTGCAATGACGACGACGACGAGAGATGCTGACAGATGCGC	573
QY	396	AGCTGGCGCATGAACACCGGGGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGC	455
DB	574	AGCTGGCGCATGAATGCTGGGGGACGCGGGAGTGTGGCGCCCAAGTATGACAGTGGT	633
QY	456	GAGATCGGCTCTCAAGTATGACAGTGGAGAGATCCCTAGGGGATACGTCCTTCAGTC	515
DB	634	GAGATCGGCTTACCAAGTACGACAGTGGTGAATCCCTCGGGGATACATCCCGTCACTC	693
QY	516	ACCMACGACGATGTCAGGAGAAATCCCTGGAGCTTGGCTGATCATCATGATGCTCC	575
DB	694	ACTAACGACGAGATTTTCGGGAGAAATCCCTGGTGTCTCCCTGACCATCATATGATGCT	753
QY	576	CCTACGGGGAACATCAGCAGACGCTGCTCCGTTTCCCTATGTGAATCATTCACCAATCCG	635
DB	754	CCTACTGGGAACATTTGGCAGCGCGCCCATTTCCCTATATGAATCATCAATCCG	813
QY	636	TCAAGGAGATTTCTCGGAGATTTGGGAATGTTGGCTGGAAAGAGAGATTTGATGGCTGG	695
DB	814	TCGAGGGAATTTCTCTGTAGCGTTGGGAATTTGCTGGAAAGAGAGGTTTGTGGCTGG	873
QY	696	AAATGAGCAGGACAGGGTGGATTTCCCATGACTAATGGCAAGCATTTGCTCCCTCT	755
DB	874	AAATGAGCAGGACAGGGAACTTCCCATGACGAATGGCAAGCATTTGCTCCCTCT	933
QY	756	GAAAGTCGGGAGCTACTGACATCGATCATCTACTGAATACAAATGGAAGAGCGCTTTA	815
DB	934	GAGGGCGGGGTGTGGTGATTTGATGATCACTCACTGATTACAAATGGAAGATGCTTTA	993
QY	816	CTGAATGATGAATCGCCAGCTCTATCTAGAAAGTCCCATGCTTCTCCCAATA	875
DB	994	TTAAACGATGAATCGCCAGCTCTATCTAGGAAGTTCACATTCCTTCTCCAGGATA	1053
QY	876	AATCCCTACGAATGGTCAATTTGCTGGGTGGTGTCTTAAGCATCTTCTCGCACTAC	935
DB	1054	AATCCATACAGATGGTCAATTTGCTGATGATTTGCTTAAGCACTCTTCTGCACTAC	1113
QY	936	CGTCTCACAATTCCTGTGCGTAATGCATACCCACTGTGCGCTTTTATCTGTTATGTGAG	995

DB	1114	CGGATCACAAATCTGTGGTAAATGATACCCACTGTGGCTTCTATCTGTTATATGTGAG	1173
QY	996	ATTGGTTTGTCTTTATCCCTGGATCTGGATCAGTTCCCGAAGTGGTTTCCAATCAACCGG	1055
DB	1174	ATCTGGTTTGTCTTTTCCCTGGATATTGGATCAGTTTCCAAAGTGGTTTCCAATCAACCGG	1233
QY	1056	GAGACCTACTCTGATAGACTGGCTTTAAGGTATGACCGGAGAAGGTGAACCGTCTCAGTTG	1115
DB	1234	GAGACTTACTCTGATAGACTCGCATTAAGTATGACCGGGAAGGTGAGCATCTCAGTTG	1293
QY	1116	GCTGCTGTTGACATATTTGTCTAGTACAGTCGACCCCTTGAAGGAGGACCATTCGTCACT	1175
DB	1294	GCTGCTGTTGACATTTTGTCTAGTACTGTCTGACCCCAATGAAGGAGGCTCCTCTGTCACT	1353
QY	1176	GCCAAACACTGTGCTATCCATTTCTGTGTTGATTATCCGTGACAAAGGCTCTTGTCTAT	1235
DB	1354	GCCAAATACCGTGTCTTCCATTTCTGCTGTGGACTATCTGTGGATGAAGTCTCTTGTCTAT	1413
QY	1236	GATCTCTGATGACGGAGCTTCAATGTGACTTTTGAACGCAFTGGCTGAGACTTCAGAGTTT	1295
DB	1414	GATCTCTGATGATGGAGCTGCTATGCTGACATTTGATGCACTAGCTGAGACTTCAGAGTTT	1473
QY	1296	GCTAGGAATGGGTACATTTTGTGAAGATGATCAATGAAACCCGAGAGCTCCCGAGTTT	1355
DB	1474	GCTAGGAATGGGTGCTATTTGTTAAAGATGACAAATGAACTAGAGCTTCCTGAAATGG	1533
QY	1356	TACTTTTGGCAGAAATTTGATTACCTGAAAGACAAAGTCCAGCTTCAATTTGTTAAAGAC	1415
DB	1534	TACTTCTCCAGAAATTTGATTACTTGAAGGACAAAGTGCACCTTCAATTTGTTAAAGAC	1593
QY	1416	CGCGGGCCCATGAAGGAGAAATATGAAGATTTTAAATCAGGATAAATGCCCTAGTTTCT	1475
DB	1594	CGCGGGCCCATGAAGGAGAAATATGAAGATTTTAAATCAGGATAAATGCCCTAGTTTCT	1653
QY	1476	AAGGCTTTGAAAGTCCCGGAGGAGATGATGATGCAAGATGCGACACCATGCGCAGGA	1535
DB	1654	AAGGCTTTGAAAGTCCCGGAGGAGATGATGATGCAAGATGCGACACCATGCGCAGGA	1713
QY	1536	AACAAATCAGGAGATCATCTCGAAATGATTGAGTTTCTTGTGTGTCACAGTGGTGGCTT	1595
DB	1714	AACAAATCAGGAGCCATCTCGAAATGATTGAGTTTCTTGTGTGTCACAGTGGTGGCTT	1773
QY	1596	GATCTGAGGCTAATGAGCTCCCGCTTTAGTTTATGTGTCTCGTGAAGAGCTGCTGGG	1655
DB	1774	GATCTGAGGCTAATGAGCTACCCGCTTTGCTGTATGTTCTCGTGAAGAGCTGCTGGA	1833
QY	1656	TTCCAGCACACAAAGAGGCTGTGCTCAATGAAATGCTGCTGCTGCTGCTGCTGCTT	1715
DB	1834	TTCCAGCATCACAAAGAGGCTGTGCTCAATGAAATGCTGCTGCTGCTGCTGCTGCTT	1893
QY	1716	ACTAATGCAATACATGTTGAAATCTTGTGATCACTACATCAACACAGCAAGCT	1775
DB	1894	ACCATGCAATACATGTTGAAATCTTGTGATCACTACATCAACACAGCAAGCT	1953
QY	1776	GTCGAGAAGCTATGTCTTCTTAATGATGCAAACTAGGTCGCGAAGTCTGTTATGTG	1835
DB	1954	CTCAGGGAAGCTATGTCTTCTTATGATCCTAACTAGGAGGAGTGTCTGCTATGTT	2013
QY	1836	CAGTTCCCAACAAAGGTTTGTGATGGATTTGATAGGATGATGATGCAACAGCAACACT	1895
DB	2014	CAGTTTCCAGAGGTTTGTGATGGATTTGATAGGATGATGATGCAACAGCAACACT	2073
QY	1896	GTCTTTTGTGATTTAACTTGGGGCTTGAACGGCTTCAAGGACAGGTTTATGTGGGA	1955
DB	2074	GTGTTTGTGATTTAACTTGGAGGCTTGTGATGGCATCCAGGACAGTTTATGTGGGC	2133
QY	1956	ACTGGTGTGTTTCAACAGAAAGCTATCTATGTTTATGAGCCCCCAATTAAGCGAAG	2015
DB	2134	ACTGGTGTGTTTCAACAGAAAGCTCTATATGTTTATGAGCCCCCAATTAAGCAAAAG	2193
QY	2016	AAGCCAGTGTCTTGGGATCACTATGTGGGGGCAAGGAGGAGGCAAGCAAGTCAAGAAA	2075
DB	2194	AAGGGTGTGTTTCTTGTCTATCTATGTTGGGAGGAGGAGGAGGCAAGCAAGTCAAGAA	2252

QY	2076	AGGAGCTCAGATAAGAAAAAGTCGAAACAGCATGTGGACAGTCTGTGTCACAGTATTCAAAT	2133
DB	2253	--GGGCTCAGACAAGAAAAGTCCAGAAAGCATGTGGACAGTCTGTGTCACAGTATTCAAAT	2310
QY	2136	CTCGAAGACATAGAGGAGGGTGTGAAAGGTCTGGGTTTGATGATGAGAAATCAGTTCTTC	2195
DB	2311	CTTGAAGATATAGAGGAGGGAGTTGAAGCGCTGGATTTGATGATGAGAAATCAGTTCTTC	2370
QY	2196	ATGTCTCAAAATCAGCTTAGAAGAAGAGATTGGCGCAGTCAGACAGCAATTTGTTGCCCTCACT	2255
DB	2371	ATGTCTCAAAATCAGCTTAGAAGAAGAGATTGGCCAATCTGCAGCTTTGTTGGCTGCACCT	2430
QY	2256	CTGATGGAATATGGTGGTGTTCCTCAGTCTCTCCACTCCAGAAATCTCTTTTGAAGAAGAGCT	2315
DB	2431	CTGATGGAATATGGTGGTGTTCCTCAGTCTCGGACTCCAGAAATCTCTCTGAAAAGAGAGCT	2490
QY	2316	ATCCATGTTCATAAGTTTGGCTATGAGGACAAGTCTGAATGGGAACTGAGATTGGTTGG	2375
DB	2491	ATCCATGTTCATAAGTTTGGCTATGAGGACAAGATTCGAATGGGAACTGAGATTGGTTGG	2550
QY	2376	ATCTATGATCTGTGCACAGAAGATATTTCTACTCGGATTCAGATGACGCAAGAGGCTGG	2435
DB	2551	ATCTATGATCTGTGCACAGAAGATATTTCTACTCGGATTCAGATGACGCAAGAGGCTGG	2610
QY	2436	CGTTCACTTATTTGCATGCCCAAGCGCCAGCTTTCAAGGATCTGCCCCCATCAATCTTT	2495
DB	2611	CGTTCGATCTACTGCATGCCCTAAGCGCGCGCTTCAAGGATCTGCCCCCATCAATCTTT	2670
QY	2496	TCAGATCGTCTGAACCAAGTCTCGGTGGGCTCTCGGTTCTGTTGAAATTTCTTTTCAGC	2555
DB	2671	TCAGACCGTCTGAACCAAGTCTCGGTGGGCTCTCGGTTCACTGGGAAATTCCTTTTCAGC	2730
QY	2556	CGGCATTTGCCCTTATCGTATGGCTACGAGGGCGCTCAAGTCTCTGGAGAGCAATTCGCT	2615
DB	2731	CGGCATTTGCCCTTATCGTACGAGGGCGCTCAAGTCTCTGGAGAGCAATTCGCTTCGCC	2790
QY	2616	TACATCAACACCACTATTTACCCACTAACCTCTCTCCGCTTCTAGTCTATTTGATATTG	2675
DB	2791	TACATCAACACCACTATTTACCCGCTCACTCCCTCCCTCCGCTCTCTATTTACTGTATCTTG	2850
QY	2676	CCGTCTATCTGTCTCTCACTGGAAAGTTCAATCATGCCAGAGATTAGCAATTTGGCCAGT	2735
DB	2851	CCGTGCCATCTGCCCTGTCTACGGGGAAGTTTCATCATCCAGAGATCAGCAACTTCGCTAGT	2910
QY	2736	ATCTGGTTTCATTCGGCTCTTCTTTCAATTTTCGCCACTGGTATCCCTGAGATGAGGTGG	2795
DB	2911	ATCTGGTTTCATTCCTCTCTTCTCTCATCTCGATTCGCCACGGGTATCTCTGAGATGAGGTGG	2970
QY	2796	AGTGGTTTGGCAATGACGATGGTGGAGAAATGAACAGTTCTGGGTCAATCGAGATATC	2855
DB	2971	AGCGGGGTGGGCATTCGACGAGTGGTGAGGAACGAGCAGTTCTGGGTCAATCGAGAGCATC	3030
QY	2856	TCCTGCACATCTGTTTGGCGTCTTTAGGGTCTTTCTGAAAGGTGGTTGCGGTAATGCAACACC	2915
DB	3031	TCGCGCCACCTCTTCGCGCTTTCTCAAGGGCTCCTCAAGGTGCTTTCGCCGCATCGACACC	3090
QY	2916	AACTTCACCTGTCACTTCAAGGCTAATGACGAAGAGGCGACTTTGCTGAGCTCTACATG	2975
DB	3091	AACTTCACCGTCACTTCGAAGGCTCGAGTGAAGACGGCGACTTCGCGGAGCTGTACATG	3150
QY	2976	TTCAAATGGAACGAGCTTCTCATCCCTCCGACGACCAATTTTCATTAATTAACATGGTTGGT	3035
DB	3151	TTCAAATGGAACGACACTTCTGATCCCGCCACACCATCTCTGATCATCAACTGTGTGGC	3210
QY	3036	GTGGTTTGTGGACCTCTTACGCCATCAACAGTGTTCACAAATGATGGGGCGCTCTTTT	3095
DB	3211	GTGTGTTGCCGGCATCTCTACGCCATCAACAGCGGGTACCACTCGTGGGGTCCGGCTCTTC	3270
QY	3096	GGGAAGCTCTCTTTTGCCTTCTGGGTGATTTGTTTCACTTATCCCATTCCTCAAGGCTCTT	3155
DB	3271	GGCAAGCTCTTCTTCGCTTCTGGGTGATGTTTCACTTATCCGTTACCTGTACCGGTTCTCAAGGGTCTC	3330

QY	3156	ATGGGCGGCAAAACCGCACACCGAATTTGTCATCGTCTGGGCTGTCCCTCGCTCTCT	3215
Db	3331	ATGGGTGGGCAAAACCGCACCGGATCGTCTGGGCGATCCCTCTGGCGTCG	3390
QY	3216	ATCTTCTCCCTTGCTGGTGGTTCGTTGATCCATTCACTACCGCTCTCGCTGGGCCAAAT	3275
Db	3391	ATCTTCTCCCTTGCTGGTGGTTCGATCGATCCGTTACCAACCGCTCACTGGCCCGAT	3450
QY	3276	ATCCAAACCTGTGGCATCAACTGCTAGGAAGTGGGA-GTTTGTAGACAGAGAA	3329
Db	3451	ACTCGAACGTTGGCATCAACTGCTAGGAGGTGGAAAGTTTGTAGAAACAGAGA	3505
RESULT 6			
LOCUS	AF200528	3745 bp	linear
DEFINITION	Zea mays cellulose synthase-4 (Cesa-4) mRNA, complete cds.		
ACCESSION	AF200528		
VERSION	AF200528.1	GI:9622879	
KEYWORDS			
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
TITLE	1 (bases 1 to 3745) Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S., Xoconostle-Cazares,B. and Delmer,D.P. A comparative analysis of the plant cellulose synthase (Cesa) gene family		
JOURNAL	Plant Physiol. 123 (4), 1313-1324 (2000)		
MEDLINE	20398328		
PUBMED	10938350		
REFERENCE	2 (bases 1 to 3745) Dhugga,K.S. and Helentjaris,T.G. Direct Submission Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA		
TITLE	Location/Qualifiers		
JOURNAL	1..3745		
MEDLINE	/organism="Zea mays"		
PUBMED	/mol_type="mRNA"		
REFERENCE	/db_xref="taxon:4577"		
AUTHORS	/note="identified and isolated from the maize genome database cDNA collection of Pioneer Hi-Bred International, Inc"		
TITLE	1..3745		
JOURNAL	/gene="Cesa-4"		
MEDLINE	320..3553		
PUBMED	/gene="Cesa-4"		
REFERENCE	/note="CESA-4"		
TITLE	/codon_start=1		
JOURNAL	/product="cellulose synthase-4"		
MEDLINE	/protein_id="AAF89964.1"		
PUBMED	/db_xref="GI:9622880"		
FEATURES			
gene			
CDS			

```

/translation="WEGDAGVKSRRGGQVCQICDGVGTTAEGDVFACDVCQGF
VQRCFEYERKDTQAPCCKYIKKHKSGPAIRBEGDDTDADSFVNLASNGDQK
QXLDERSWRNMGSGDGRPKYDGSBIGHTKDYSDGIPRGYIPSVNTSLSGSIP
GASPDHMMHPTGNIKGRAPFYVNHSPNSPFSGGINVAKWERKVGDMKQDAGT
IPMTGTSTAPSGEVRGVDIDASTYDNDENDALLINTROPGLKVPSPSRINPYMW
IVLRIVLISILFHYIRITPNRVANPLMLSLICEWFAISLWIDQFKPKNRYEVL
DLRALYDREGSPQLAAVDIFSVTVDMPKPELVTANTVLISLAVDVPDKVCSYV
DDGAAMLTDLAALTESFAKWWPVFKYKINBPRAEPFESQIDIKDKVHSPFKVD
RMAKREYEFKRVNGLVAKAQVPEEGWIMQDGTPEWNTDHPGMIQVPLGHSG
GLDTSEGNPLFVDSYREKRGCPQHHKAGANALVRUSAVLTNGOYMLNLCDCHVIN
NSKALREAMCLMVDNLNRCGVCPQBPQDQFIDNRDNYANRNTFFDINLRGLQ
GPVYVQTCVENETALGXVERPIKQKGFSLSCGRKASKSGKSGKSGKQKRD
SPVVPNVEDIBEGVEGAGFDEBSLIMSOMLSRGLQSGAFAPVLTMEYGVGPOA
TPBSLLKEAHVISCYEDETBWEITGIMYGSVETDILTFPMHARGWRSYCMKPR
FAPKGSAPINSDRNLVRWALGSEVDFHSCHLVPVWGYGKGLERFAVNTIYV
TISLPIILYCLIPALTKGFIITPEISNFAISWISLFIISFATIGLEPMWTSVGI
DEWNRNDFWIGTSJSAHLFAVQGLLVKAGIDFTNVTSKASDBDQFIAELMYEKV
TTLTLPPTTILINJUVAGVAYSINQSGWGLFGLFAFWYHLVPTPKGLM

```

GRQNRPTIVVWAILLASIPSLWVRIDPFTTRVTRVTPDTQTCGINC"

ORIGIN

Query Match 70.9%; Score 2572.2; DB 8; Length 3745;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2;

QY	4	CGAGGAACCCCGCTCCAGCTCTGTCTGTCGGTGGGTTGGATCGCTCTCGCGGCCCATGG	63
DB	264	CCAGGCCCCAGGCTCCAGGCCAGCTCCCTCGACGTTTCTGGCGAGCTCGCTGGCCATGG	323
QY	64	ACGGCCACGGGACGCCCTGAGCTCCGGGAGGACACGGGCGCGGGAGCGTGTGCCAGATCT	123
DB	324	AGGGCCACGGGACGGGCTGAGCTCCGGGAGGCGCGTGTGGCGGACAGGTGTGCCAGATCT	383
QY	124	GCSCCGACGGCTGGGACCCACGCTTGGACGGCGACGCTTTCACCGCTCGGACGTGTGCC	183
DB	384	CGCGGACGCGGTGGGACCCACGCGGAGGGGACGCTCTTCGCGCGCTCGGACGCTGTGG	443
QY	184	GCTTCCGGTCTCGCGCCCTGTCTAGCAGCAGCGCGCAGGAGGGCACCCAGGCTGTGCC	243
DB	444	GGTTCCGGTGTCCGCGCCCTGTCTAGCAGTACGAGCGCAAGGACGGCACCGAGGCTGTGCC	503
QY	244	TCCAGTGCAGACCAAGTACAAGCGCCACAGAGGGAGCCCGAGCGATCCGCGGGGAGGAAG	303
DB	504	CCAGTGCAGACCAAGTACAAGCGCCACAGGAGGGAGCCCGAGCGATCCGCGGGGAGGAAG	563
QY	304	CGCAGCACACTGATGCCGATGATGTAGTGACTTCAACTACCTGCACTGTGGCACTGAGG	363
DB	564	GAGACGACACTGATGCCGA-----TAGCGACTTCAATTACCTTGGCATCTGGCAATGAGG	617
QY	364	ACCAGAGCAGAGATTTCTGACAGGATCGGAGCTGGCGCATCAACACCGGGGGCAGTG	423
DB	618	ACCAGAGCAGAGATTTCCGACAGAAATGGCGAGCTGGCGCATGAACGTTGGGGGCGAGCG	677
QY	424	GCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCGCTCTCCAAGTATGACAGTG	483
DB	678	GGGATGTTGGTGGCCCCCAAGTATGACAGTGGCGAGATCGGCGCTTACCAAGTATGACAGTG	737
QY	484	GAGAGATCCCTAGGGGATAGCTCCCTTCAAGTACCAACAGCGAGATGTACGAGAAATCC	543
DB	738	CGGAGATTCCTCGGGGATACATCCATCAGTCACTAACAGCCAGATCTCAGGAGAAATCC	797
QY	544	CTGAGCTTCGCTGATCATCATGATGTCCCTTACGTCAGGGAACATCAGCAGAGCTGCTC	603
DB	798	CTGGTGTCTCCCTGACCAATCATATGATGTCCCACTGGGAACATTTGGCAAGCGTCTC	857
QY	604	CGTTCCCTATGTGAATCATTCACCAATCCGTCAAGGGAGTTCTCCGGCAGTATTGGGA	663
DB	858	CATTTCCCTATGTGAACCATTCGCCAATCCGTCAAGGGAGTTCTCTGGTAGCATTTGGGA	917
QY	664	ATGTTGCTGGAAGAGAGAGTGTGATGGCTGGAAATGAAGCAGGACAGAGGTGCGATTG	723
DB	918	ATGTTGCTGGAAGAGAGGTTGTATGGCTGGAAATGAAGCAGGACAGAGGTGCGATTG	977
QY	724	CCATGACTATGGGACAGCATTTCTCCCTCTGAGGTGCGGCGAGTACTCAGATCGATG	783
DB	978	CCATGACGAATGGGACAGCATTTCTCCCTCTGAGGTGCGGCGAGTACTCAGATCGATG	1037
QY	784	CATCTACTGAATACAACTAGGAAGACGCTTTACTGAATGATGAATCTGCCAGGCTCTAT	843
DB	1038	CATCAACTGATTAACAATGAGAGATGCTTTATTGAACGACGAACTCGACAGCTCTAT	1097
QY	844	CTAGAAAGTCCCATTCCTTCCCAAAATAAATCCCTCAGAAATGATGATGTTCTGCG	903
DB	1098	CTAGAAAGTCCCATTCCTTCCCTCCAGGATAAATCCATACAGGATGATGATGTTCTGCG	1157
QY	904	GTTGGTGTCTAAGCATCTTCGCACTACCGCTCACAATCTCCTGCGTGAATGAT	963
DB	1158	GATTGATGTTCTAAGCATCTTCGCACTACCGTATCAAAATCTGTTGGCGAATGAT	1217
QY	964	ACCCACTGTGGCTTTATCTGTATATGTGAGATTTGTTGCTTTATCTCTGGATCTGG	1023

DB	1218	ACCATTATGGCTTCTATCTGTATATGTAGATCTGGTTTGTCTCTTTCGTGTGATATGG	1277
QY	1024	ATCAGTTCGCCGAAGTGGTTTCCAAATCAACGGGAGACCTTACCTTGTATAGACTGGCTTTAA	1083
DB	1278	ATCAGTTCCTTAAGTGGTTTCCAAATCAACGGGAGACCTTACCTTGTATAGACTGGCTTTAA	1337
QY	1084	GGTATGACCGGAGAAGGTGAACCGTCTCAGTTGGTGTCTGTGTGACATATTTGTGACGTACAG	1143
DB	1338	GGTATGACCGGAGAAGGTGAGCCATCTCAGTTGGTGTCTGTGTGACATATTTGTGACGTACAG	1397
QY	1144	TCGACCCCTTGAAGAGGCCACCTATGCTCACTGCCAACACTGTGCTATCCATTTCTGTCTG	1203
DB	1398	TCGACCCCAATGAAGGAGCCCTCTCTTGTCACTGCCAAATACCGTGTCTATCAITTTCTGTG	1457
QY	1204	TTGATTATCCGCTGGGACAAAGGTCTCTTGTATGTATCTATGATGACGAGGTTCATATCTGA	1263
DB	1458	TGATTTATCCCTGTGGATTAAGTCTCTTGTCTATGTATCTGATGATGAGCTGCGATCTGA	1517
QY	1264	CTTTTGACGCAATTTGGCTGAGACTTTCAGAGTTTCTAGGAAAATGGGTACCCATTTGTGAAGA	1323
DB	1518	CATTTGATGCATCTAGCTGAGACTTTCAGAGTTTCTAGAAAATGGGTACCCATTTGTGAAGA	1577
QY	1324	AGTATGACATTTGAACCCAGAGCTCCGAGTTTACTTTTCCAGAAAAATTTGATTAATCTGA	1383
DB	1578	AGTATGACATTTGAACCTAGAGCTCTGAAATGGTACTTCTCCAGAAAAATTTGATTAATCTGA	1637
QY	1384	AAGACAAAGTCCAGCCCTTCATTTGTTAAAGACCGCGGGCCATGAAGAGAGAAATTAAG	1443
DB	1638	AGGACAAAGTGCACCTTCATTTGTTAAAGACCGCGGGCCATGAAGAGAGAAATTAAG	1697
QY	1444	AATTTAAATTCAGATAAATGCCCTAGTTTCTTAAGGCAITTTGAAAGTCCCGAGGAGGAT	1503
DB	1698	AATTTAAATTCAGATAAATGGCTTGTCTTAAGGCACAGAAAGTTCTCTGAGGAGGAT	1757
QY	1504	GGATCATGCAAGATGGCACACCATGCGCCAGGAACAAATACAGGGATCATCTCTGGAATGA	1563
DB	1758	GGATCATGCAAGATGGCACACCATGCGCCAGGAACAAATACCMGGACCATCTCTGGAATGA	1817
QY	1564	TTCAGGTTTCTTGGTTCAGGTGGTGGCTTGTATCTAGGGGTAAATGAGTCCCGGTT	1623
DB	1818	TTCAGGTTTCTTGGTTCAGGTGGTGGCTTGTATCTAGGGGTAAATGAGTCCCGGTT	1877
QY	1624	TAGTTTATGTCTCTGTAAGAAAGCGTCTCGGTTCAGACCAACAAAGGCTGGTGCCA	1683
DB	1878	TGGTCTATGTTCTCTGTAAGAAAGCGTCTCGGTTCAGACCAACAAAGGCTGGTGCCA	1937
QY	1684	TGAATGCCCTTGTCTGTTCTCAGCTGTCTTACTAATGACAAATCATGTTGAATCTTG	1743
DB	1938	TGAATGCCCTTGTCTGTTCTCAGCTGTCTTACTAATGACAAATCATGTTGAATCTTG	1997
QY	1744	ATTGTGATCACTACATCAACAAAGCAAGGCTGTCCGAGAAGCTATGTGTTCTTAATGG	1803
DB	1998	ATTGTGATCACTACATTAACAAAGTAAGGCTCTCAGGAGGCTATGTGTTCTTAATGG	2057
QY	1804	ATCCAAACCTAGGTCCGCAAGTCTGTTATGTGCAAGTCCCAACAAAGGTTGATGGGATTG	1863
DB	2058	ACCCTAACCTAGGAAGGAGTGTCTGTCTACGTCCAGTTTCCCGAGAGTTCCGATGCAATG	2117
QY	1864	ATAGGAATGATCGATATGCAACAAAGCAAGCACTGTCTTTTGTATTAATTAACCTTGAAGG	1923
DB	2118	ACAGGAATGATCGATATGCAACAAAGCAAGCACTGTCTTTTGTATTAATTAACCTTGAAGG	2177
QY	1924	TTGACGGCATTCAGAGGACCGAGTTTATGTGGGAATCTGTGTTGTTTCAACAGAAACAGCTA	1983
DB	2178	TTGATGGCATTCAGAGGACCGAGTTTATGTGGGAATCTGTGTTTCAACAGAAACAGCTC	2237
QY	1984	TCTATGTTATGACCCCGCAATTAAGGGAAAGAGCCAGGTTTCTTGGCATCACTATG	2043
DB	2238	TATATGTTATGACCCCGCAATTAAGGAAAGAGGAGGTTTCTTGTTCATCACTATG	2297
QY	2044	GGGCAAGAAAGCAAGCAAGCAAGTCAAGAAAAAGGAGCTCAGATAAGAAAAAGTTCGAACA	2103
DB	2298	CGGTAGGAAGAGCAAGCAAGCAAGTCAAGAAAAAGGAGCTCAGATAAGAAAAAGTTCGAACA	2354

Qy	2104	AGCATCTGGACAGTCTTGTTCCAGTATTCAATCTCGAAGACATAGAGGAGGGTCTTGAAG	2163
Db	2355	AGCATGTGGACAGTCTTGTCGCCAGTATTCAACCTTGAAGATATAGAGGAGGGAGCTTGAAG	2414
Qy	2164	GTGCTGGTGTGATGATGAGAAATCAGTTCTTCATGTCTCAATCAGCTTTAGAGAAGAGAT	2223
Db	2415	GGCTGGATTGACGACGAGAAATCACTTCTTATGTCTCAATGAGGCTTGGAGAAGAGAT	2474
Qy	2224	TTGGCCAGTCAGCAGCATTTGTTGCCCTCCACTCTGATGGAATATGGTGGTGTTCCTCAGT	2283
Db	2475	TTGGCCAGTCGCGCAGCGTTTGTGGCTTCCACTCTGATGGAGTATGGTGGTGTTCCTCAGT	2534
Qy	2284	CCCTCACTCCAGATCTCTTTTGAAGAAGCTATCCATGTCATTAAGTTGTGGCTATCAGG	2343
Db	2535	CCGCACTCCGGAGTCTTCTCTGAAGAAGAGCTATCCATGTTTATAAGCTGTGGCTATCAGG	2594
Qy	2344	ACAAGTCTGAATGGGGAACCTGAGATTGGTTGGATCTATGGATCTGTGCACAGAAGATATTC	2403
Db	2595	ACAAGACTGAATGGGGAACCTGAGATCGGTGGATCTACGGTTCGTGTACAGAAGACATTC	2654
Qy	2404	TTACTGGAATCAAGATCGACCAAGAGCGCTTGGCGTTTCAGTCTATPAGCATGCCCAAGCGCC	2463
Db	2655	TCACCGGATTCAGATGACGCGCGAGCGCTGGCGGTGCACTACTGTCATGCCCAAGCGGC	2714
Qy	2464	CAGCTTTCAGGGATCTGCCCCCATCAATCTTTTCAGATCGCTGTGAACCAAGTGTGCGGT	2523
Db	2715	CAGCTTTCAGGGGTCTGCCCCCATCAATCTTTCGGACCGTCTGAACCAAGTGTCTCGGT	2774
Qy	2524	GGGCTCTCGGTTCTGTGTGAAATTCCTTTTCAGCCGGCATATGCCCTTATGTTATGGCTACG	2583
Db	2775	GGGCTCTGGGTCTGGTGGAGATCCTCTTCAGCGGCACTGCCCTGTGGTACGCTACG	2834
Qy	2584	GAGGGCCCTCAAGTTCCTCTGAGAGATTGGTTTACATCAACACACCAATTTACCCACTAA	2643
Db	2835	GAGGGCGGCTCAAGTTCCTGAGAGATTGCGGTACATCAACACCAATCTACCGCTCA	2894
Qy	2644	CCCTCTCCCGCTTCTAGTCTATTGTATATTGCTGCTATCTGCTCTGCTCACTGGAAGT	2703
Db	2895	CGTCCATCCCGCTTCTCATCTACTGCACTCTGCCCGGCATCTGCTGCTCACCGGAAAGT	2954
Qy	2704	TCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTCAATGCGCTCTTCTTTCAA	2763
Db	2955	TCATCATTCAGAGATCAGCAACTTCGCCAGCATCTGGTTCACTCCCTCTTCATCTCGA	3014
Qy	2764	TTTTCGCCACTGGTATCTTGTAGATGAGGTGGAGTGGTGTGGCATTTGACGATGGTGGGA	2823
Db	3015	TCCTTGGCCACGGGCATCTCTGGAGATGAGGTGGAGCGGGGTGGGCATCGACGAGTGTGGGA	3074
Qy	2824	GGAAATGAACAGTTCTGGGTCAATGGAGGTATCTCTGCACATCTGTTTGGCGTCTTTCAGG	2883
Db	3075	GGAAACGAGCAATTTCTGGGTGATCTGGGGGCACTCTCCGGGCACTCTTCGCCGTGTTCAGG	3134
Qy	2884	GTCTTCTGAAGGTGCTTTCGCGGTATCGACACCAACTTTCACCTGTCACTCAAGAGCTAATG	2943
Db	3135	GCCTGTCTCAAGTGTCTGCCCGGCATCGACACCAACTTTCACCGTCACTCAAGGCCTCG	3194
Qy	2944	ACGAGNAGCGGACTTTGCTGAGCTCTACATGTTCAAGTGGACGACGCTTCTCATCCCTC	3003
Db	3195	ACGAGGACGGCACTTCGCGGAGCTGTACATGTTTCAAGTGGAGGAGCGTCTCTGATCCCGC	3254
Qy	3004	CGACGACATTTTGATCATTTAAATGTTGGTGTGGTGTGGCACCTCTCAAGCCATCA	3063
Db	3255	CCACCACCATCTGATCATCAACTGGTCTGGGTCTGGTCTGGCACCTCTCAAGCCATCA	3314
Qy	3064	ACAGTGTTTACCAATCATATGGGGGCGCTCTTTTGGGAAGCTCTTCTTTTGGCTTCTGGGTGA	3123
Db	3315	ACAGCGGATACCAAGTCGTGGGGCCGCTCTTCGCGAAGCTCTTCTTCGCTTCTGGGTCA	3374
Qy	3124	TTGTTTCACTTATACCAATTCCTCAAGGGTCTTATGGGACGGCAAAACCGCACACCGACGA	3183
Db	3375	TCGTGCACTGTATCCCGTCTCTCAAGGGCTCTATGGCAGCGAGAACCGCACCCGACCA	3434

Qy	3184	TTGTCATCGTCTGGCGTGTCTCTCGTCTTCTATCTTCTCTCTGCTGTGGTCTCGTGTG	3243
Db	3435	TCGTCGTCTGTGGCCATCTGTCGTGCGTCCATCTTCTCTCTGCTGTGGTTCGCATCG	3494
Qy	3244	ATCCATTCATCACTACCCGCTCGCTGGCCCAATATCCAAACCTGTGGCATCAACTCTAGG	3303
Db	3495	ACCCCTTCACACCCCGCTCAGTGGCCGGATACCCAGAGTGTGGCATCAACTGCTAGG	3554
Qy	3304	AAAGTGGGAGTTT	3316
Db	3555	GAAGTGAAGGTT	3567

RESULT 7
 AK069196
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AK069196 4282 bp mRNA linear PLN 24-JUL-2003
 Oryza sativa (japonica cultivar-group) cDNA clone:J023003G18, full insert sequence.
 AK069196
 FLI cDNA; CAP trapper.
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
 AUTHORS
 1 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,M., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group.; Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Narikawa,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narioka,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN; Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
 Science 301 (5631), 376-379 (2003)
 22752273
 12863764

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

2 (bases 1 to 4282)
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Oeato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
 Direct Submission
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of

TITLE
 JOURNAL

Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica rice.

COMMENT

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, S., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, K., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

```
1. 4282
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023003G18"
```

FEATURES

source

ORIGIN

Query Match 65.1%; Score 2360.8; DB 8; Length 4282;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 2690; Conservative 0; Mismatches 527; Indels 3; Gaps 1;

```
QY 529 TGTGAGAGAAATCCCTGGAGCTTCGCTGATCATCATCATGATGTCCCTACGGGAACA 588
    |||
Db 726 TCTCAGTGAATTCCTGGAGCGTCCCTGATCATATGATGTCTCCGTTGGGAACATTG 785
QY 589 TCAGCAGAGTGCTCGTTTCCCTATGTGAATTCACCAATCCGTCAAGGAGTTCT 648
    |||
Db 786 GCAGAGCTGGGCATCCATTTCCCTATGTGAACCAATTCACCAACCCATCAAGGAGTTT 845
QY 649 CGGGCAGTAGTGGGAATGCTGCTGAAAGAGAGAGTGTGATGGCTGCAAAATGAAGCAGG 708
    |||
Db 846 CTGGTAGCCCTTGCAATGTTGATGAAAGAGAGAGTAGATGGCTGCAAAATGA---AGG 902
QY 709 ACAAGGGTCCGATTCCTCATGACTTAATGGGACAGCAATTCCTCTCTGAAGTTCGGGCGAG 768
    |||
Db 903 ATGAAGGTGCAATTCCTCATGGCTAATGGTACTTAGCATTTGCTCTTCAGAAGGTCTGGGG 962
QY 769 CTACTGACATCGATGATCTACTGATACACATGAAAGAGCGTTTACTGAAATGATGAAA 828
    |||
Db 963 TTGGTGACATTTGATGATCTACTGACTATAACATGAGGAGCGCTTACTGAATGATGAAA 1022
QY 829 CTCGCCAGCTCTATCTAGAAAAGTCCCATTTGCTTCTCCAAAATAAATCCCTACAGAA 888
    |||
Db 1023 CAGCCAACTCTATCGAAGAAAGTCCCAATTTCTGTCATCCAGATTAATCCGTACAGGA 1082
QY 889 TGTCTATTGTTCTGCGGTTGGTTTCTAAGCATCTTCCTGCACTACCGTCTCACAAATC 948
    |||
Db 1083 TGGTTATTGTTCTCGGATTGTTGCTATGTATATTTCTGCACTACCGTATCACAAATC 1142
QY 949 CTGTGGTAAATGATACCCACTGTGCTTTATCTGTATATGATGAGATTTGTTTGCCT 1008
    |||
Db 1143 CTGTGGGTAATGATACCCGCTATGTTGCTCTCTCATATGATGAGATTTGTTTGCCT 1202
QY 1009 TATCTCGTACTGATGATTCCTCCGAAGTGGTTTCCAAATCAACCGGAGACCTACCTTG 1068
    |||
Db 1203 TATCTCGTACTGATGATTCCTCCGAAGTGGTTTCCAAATCAACCGTGAATCTACTGG 1262
QY 1069 ATAGACTGCTTTAAGTATGACCGGAGAGTGAACCGTCTCAGTTGGCTGTGTTGACA 1128
    |||
Db 1263 ATAGGCTGGCTTTAAGTATGACCGGAGAGTGAACCGTCTCAGTTGGCTGTGTTGACA 1322
QY 1129 TATTTGTCAGTACAGTCGACCCCTTGAAGAGAGCACTATCGTCACTGCCCAACACTGTGC 1188
    |||
Db 1323 TTTTGTGTCAGTACTGGATTCCTATGAAAGAACTCTCTTGTCACTGCCCAACTGTGC 1382
QY 1189 TATCCATTTCTGCTGTTGATTATCCCGTGACAAAGTCTCTGCTATGTATCTGATGACG 1248
    |||
Db 1383 TCTCCATCTCTGCTGATGATTAACCTGTTGATAGGTATCTCTGTAATGCTGTGATGATG 1442
QY 1249 GAGCTTCAATGCTGACTTTTGAAGCATTTGGCTGAGACTTCAGAGTTTGTAGGAAATGGG 1308
    |||
Db 1443 GAGCTGCAATGTTAAACATTTGATGCGCTTCTGCTGAGACTTCAGAAATTTGCTAGAAAATGGG 1502
QY 1309 TACATTTGTGAAGAGATGACATTTGAACCCAGAGCTCCGAGTTTACTTTTGCCAGA 1368
    |||
Db 1503 TACCGTTCTGTAAAAAGTACAGCATAGAACCCAGAGCTCCAGAGTGTGATCTTTGACAGA 1562
QY 1369 AAATTGATTACCTGAAAGACAAAGTCCAGGCTTCATTGTTTAAAGACCGCGGCGCATGA 1428
    |||
Db 1563 AAATTGATTACTTGAAGACAAAGTTCAGGCTTTCTTTGTTTAAAGATCGTGTGCGCATGA 1622
QY 1429 AGAGAGATATGAAGAAATTTAAATACAGGATAAATGCCCTCAGTTTCTAAAGCATTTGAAG 1488
    |||
Db 1623 AGAGGGAATATGAAGAAATTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAG 1682
QY 1489 TCCCGGAGGAGATGATCATGACAGATGCGACACCATGCCAGGAAACAAATACCAAGG 1548
    |||
Db 1683 TTCTGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1742
QY 1549 ATCATCTCTGGAATGATTCAGGTTTCTCTGCTCAGTGTGGCTTGTGATCTAGTGGGTTA 1608
    |||
Db 1743 ACCATCTCTGGAATGATTCAGGTTTCTCTGCTCAGTGTGGGCTTGTGATCTAGTGGGTTA 1802
QY 1609 ATGAGCTCCCGGCTTTAGTTTATGTTCTCGTGAAGAGCGTCTCGGTTTCCAGCACCA 1668
```


Db	32	ATGCTGCCAGATCTCGCGGACGGCTGGGCAACGGCGGAGACGGGAGCTCTTCAACG	91	1109	TCTCCATCCTTGCTGTGGATTACCCCTGTTGATAAGGTATCTTGTCTATGTGTCTGTGATG	1168
Qy	169	CTTGCACTGTCGCGCTTCCCGGTTCCCGGTTGTCGCGCCCTTGCTACGAGACAGAGCGCAAGAGG	228	1249	GAGCTTCAATGCTGACTTTTGGCGCATTCGGCTGAGACTTCAGAGTTTGTAGGAAATGG	1308
Db	92	CTGCGACGCTCTGCGGGTTCCCGGTTCCCGGTTGTCGCGCATGCTACGAGTACGAGCGCAAGGACG	151	1169	GAGCTGCAATGTTAAACATTTGATGCGCTTGTCTGAGACTTTCAGAAATTTGCTAGAAAATGG	1228
Qy	229	GCACCCAGGCGCTCTGAGTGCAGACCAAGTACAAAGCGCCACAGAGGAGCGCCAGCGA	288	1309	TACCATTCTGTAAGAAGTATGACATTTGAACCCAGAGCTCCCGAGTTTACTTTTGCAGA	1368
Db	152	GCAGCCAGGCGTCCCGCGAGTGCAGCAAGTACAAAGTACAAAGCGCCACAAAGGAGGCGCCCGCA	211	1229	TACCCTTCTGTAAGAAGTACAGATAGAACCCAGAGCTCCAGAGTGGTACTTTGACAGA	1288
Qy	289	TCCGCGGGGAGGAGCGACGACACTGATGCCGATGATGCTAGTGTGACTTCAACTACCCCTG	348	1369	AAATGATTAACCTGAAAGCAAAAGTCCAGCTTTCATTGTTAAAGACCGCGCGGCGCATGA	1428
Db	212	TACTTGGGGATGAAGGATGATGTTGATCGGATGATGCTAGTGTGAACTATCCAA	271	1289	AAATTCATTTACTTTGAAAGCAAAAGTTCAGGCTTCTTTTGTAAAGATCGTCGTGCCATGA	1348
Qy	349	CATCTGGGCTGAGGACCGAGACAGAAAGTCTGACAGGATGCGCAGCTGCGCATGA	408	1429	AGAGAGAATATGAAGAAATTTAAATCAGGATAAAATGCCCTAGTTTCTTAAGGCAATGAAG	1488
Db	272	CATCCGGCAACACGAGCACCAAGCAAGATTCCCGAGAGGATGCTCACCTGCGCATGA	331	1349	AGAGGGAATATGAAGAAATTTAAAGTTCGTTGTTAATGCTCTCGTCGCGCAAGGACAGAAAG	1408
Qy	409	ACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCTCT	468	1489	TCCCGGAGGAAGGATGGATCATGCAAGATGGCAACCATGGCCAGGAAACAAATACCAGGG	1548
Db	332	ACTCTGGGAGGATGATGATATTTGTCATTTCTAAGTATGATGATGATGATGATGATGATG	391	1409	TTCTCTGAGGAAGGATGATGATGCAAGATGGCACACCTTGGCTGGTAAACATACCAGAG	1468
Qy	469	CCAAGTATGACAGTGGAGAGATCCCTAGGGGATACCTTCCCTTCAGTACCAACAGCCAGA	528	1549	ATCATCCTCGAATGATTCAGGTTTTCCTTGGTCCAGTGGTGGCTTGTATCTGATCTGAGGTA	1608
Db	392	CCAAGTATGACAGTGGTGAATCCCTCGCATATATATCCATCGCTCACTCACAGCCAGA	451	1469	ACCATCCTGGAATGATTCAGGTTTTCCTTGGTCCATAGTGGAGGCTTGTATACCAGAGGCA	1528
Qy	529	TGTCAGAGAAATCCCTGAGCTTCCGCTGATCATCATGATGTCCTTACCGCTACCGGGAACA	588	1609	ATGAGCTCCCGCTTGTAGTTTATGCTCTCGTAAAGCGTCTCGGTTCCAGCACCCACA	1668
Db	452	TCTCAGGTGAATTCCTGAGCGTCCCTGATCATATGATGCTCCCGTTGGGAAACATTTG	511	1529	ATGAGCTTCTCTGTTAGTCTATGTCCTCGTGAGAAACGTCTCTGGATTTCCACACCATTA	1588
Qy	589	TCAGAGAGCTGCTCGGTTTCCCTATGTAATCATTCACCAATCCGTCAGGAGGATTTCT	648	1669	AGAAGCTGGTGCATGAATGCCCTTGTTCGTGTCTCAGCTGTCTTACTTAATGGAACAAT	1728
Db	512	GCAGAGTGGGATCCATTTCCCTATGTAACCATTTCAACCAACCATCAAGGAGGATTTT	571	1589	AAAAGCTGGTGTATGAATGCATTTGTTTCGTGTATCAGCTGTCTTACTTAATGGAACAAT	1648
Qy	649	CCGCGAGTATTTGGGAATTTGCTCGGAAAGAGAGAGTTGATGCTGGAAATGAAGCAGG	708	1729	ACATGTTGAATCTTGTGATCACTACATCAACACAGCAGAGCTGTCCGAGAGACTTA	1788
Db	572	CTGGTAGCTTGGCAATTTGTCATGTAAGAGAGAGTAGATGCTGGAAATGA--AGG	628	1649	ACTTGTGAACTCTGATCTGTGATCACTACATCAACATAGCAAAAGCTCTCCGAGAGGCTA	1708
Qy	709	ACAAAGGTGGATCCAGTACTAATGGAAGCAAGCATTTCTCTGTAAGGTGCGGCGAG	768	1789	TGTCCTTCTTAATGGAATCAAACTAGTCCGCAAGTCTGTTATGTGCAAGTCTCCCAAAA	1848
Db	629	ATAAAGGTCAATTCCTGAGCTTAAATGTTACTAGCATTTGCTTCTCAGAAAGTCTGGGG	688	1709	TGTCCTTCTTAATGGAATCAAACTAGTCCGCAAGTCTGTTATGTGCAAGTCTCCCAAAA	1848
Qy	769	CTACTGACATCGATGATCTACTGTAATACAAATGGAAGAGCTTTACTCAATGATGAAA	828	1849	GCTTTCATGGGATTTGATAGGAATGATCGATATGCAAAACAGGAACTGTCTTTTTCGATA	1908
Db	689	TTGGTGACATGATGATCTACTGATCTAATGAGAGAGCGCTTACTGATGATGATGAAA	748	1769	GCTTTCATGGGATTTGATAGGAATGATCGATATGCAAAACAGGAACTGTCTTTTTCGATA	1828
Qy	829	CTGCCAGGCTCTATCTAGAAAAGTCCCAATTCCTCCAAATAAATCCCTACAGAA	888	1909	TAACTTGAAGGCGCTTGAAGGATTCAGGACCTGTTATGTGGAACTGTGTTGTTT	1968
Db	749	CACGCCAACCTCTATCGAGAAAAGTCCCAATTTCTGTCATCCAGAAATAATCCGTACAGGA	808	1829	TAAATTTGAGAGTCTCGATGGTCTCCAAAGGACCAAGTTTATGTGGAAACCGTTGTGT	1888
Qy	889	TGTCATTTGTTGCGGTTGGTTGTTCTAAGCATCTCTGCACTACCGTCTCACAAATC	948	1969	TCAACAGAACAGCTATCTATGTTTATGAGCCCAATTAAGGCGAAAGACCGAGTTTCT	2028
Db	809	TGCTTATGTCCTCCGATGATTTGCTCTCTGTCATATGATGATGATGATGATGATGATG	868	1889	TCAACAGAACAGCTCTTATGTTTATGAAACCCCAATTAAGCAGAGAGGCGCAGTTACT	1948
Qy	949	CTGTCGTAATGATACATACCACTGCTGGCTTTTATCTGTTATGATGATGATGATGATG	1008	2029	TGGCATCCTATGTTGGGGCAAGAAAGCAAGCAAGTCAAAAGAAAGAGGCTCAGATA	2088
Db	869	CTGTCGTAATGATACATACCACTGCTGGCTTTTATCTGTTATGATGATGATGATGATG	928	1949	TCTCTTCGCTTTTGTGGGGCAAGAAAGACAAAAAGTCCAAAGGAGAGAGCAGGAAA	2008
Qy	1009	TATCTGGATACATGATGATTCGGAAGTGGTTTCAATCAACCGGAGACCTACCTTG	1068	2089	AGAAAAGTTCGAACAGCATGTGGAAGTTCCTGTTCCAGTATTCATCTCGAAGACATAG	2148
Db	929	TATCTGGATACATGATGATTCGGAAGTGGTTTCAATCAACCGGAGACCTACCTTG	988	2009	AGAAAAGTTCGAACAGCATGTGGAAGTTCCTGTTCCAGTATTCATCTCGAAGACATAG	2068
Qy	1069	ATAGATGCTGTTAAGGTATGACGAGAGGTTGAACCGTCTCAGTTGGCTGCTGTTGACA	1128	2149	AGAGAGGTTGTAAGGTTGCTGGGTTTGTATGATGAGAAATCAGTTCCTCATCTCAAAATGA	2208
Db	989	ATAGGCTGGCTTTAAGGTATGACGAGAGGTTGAACCGTCTCAGTTGGCTGCTGTTGACA	1048	2069	AGGAAGGATTTGAAGTTCTGGATTTGATGATGAGAAATCACTACTGATCTCAAAATGA	2128
Qy	1129	TATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1188	2209	GCTTAGAAGAGATTTGGCCAGTACGAGCATTTGTTGCTCCACTCTCTGATGGAATATG	2268
Db	1049	TATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1108	2129	GCTTAGAAGAGATTTGGTCAATCTAGTGTCTGTCGAGTTTATCTTGAAGATATAG	2188
Qy	1189	TATCCATTTGCTGTTGATTTCCCGTGAGCAAGGTTCTTGTGATGATCTGATGATG	1248	2269	GTGGTCTCTCAGTCTCCACTCCAGATCTCTTTTGAAGAGAGCTATCCATGTCTATAA	2328
				2189	GTGGTCTCTCAGTCTCCACTCCAGATCTCTTTTGAAGAGAGGCTATCTGTTATCA	2248

QY	2329	GTGTGGCTATGAGGACAAGTCTCAATGGGAACTGAGATTGGTTGGATCTATGGATCTG	2388
Db	2249	GCTGTGGCTATGAAGACAAAAGTGACTGGGAACTGAGATTGGATGGATCTATGGTTCTG	2308
QY	2389	TCACAGAAGATATTTCTTACTCGATTCAAGATGACGGAAGAGGTCGGTTCAGTCTATT	2448
Db	2309	TTACAGAAGATATTTCTCACTGGATTCAAGATGATGACGCTGGTGGCGGTCAATCTACT	2368
QY	2449	GCATGCCAAGCGCCAGCTTTTCAAGGGATCTGCCCCATCAATCTTTTCAGATCGTCTGA	2508
Db	2369	GCATGCCCTAAGCGACCGCAATTCAGGGGTCTGCCCAATTAATCTTTCCGATCGTCTGA	2428
QY	2509	ACCAAGTGCTGCGTGGGCTCTCGGTTCTGTTGAAATTTCTTTTCAGCGGCATTGGCCCT	2568
Db	2429	ACCAGGTGCTTCGGTGGGCTCTTGGTTCTCTGGAAATTTCTTTTAGCAGCAATGTCCCA	2488
QY	2569	TATGGTATGCTACGAGGGGGCTCAAGTTCTGAGAGATTGCGTTAATCAACACCA	2628
Db	2489	TATGGTATGGGTACGGAGGACGGCTTAAGTTCTCGAGAGATTGCTTTACATCAACACCA	2548
QY	2629	CCATTTACCCACATACTCTCTCCGGTTCAGTCTATCTATATTGCGTCTATCTGTC	2688
Db	2549	CTATTTATCCACTCACATCGAATTCACCTCTCTTGTACTGCATATTGCCAGCTATCTGTC	2608
QY	2689	TGCTCACTGGAAAGTTTCATCATGCCAGAGATTAGCAACTTTGGCCAGTATCTGGTTCAATTG	2748
Db	2609	TTCTCACTGGGAAGTTTCATCATACCGAGATTAGCAACTTTGCCAGTATTTGGTTATCT	2668
QY	2749	CGCTCTTCCTTTCAATTTTCGCCACTGGTATCTCTGAGATGAGGTGAGTGGTTGGCA	2808
Db	2669	CTCTCTTTCTGTCAATCTTTTGCTACTGGTATCTCTGAGATGAGGTGAGTGGTTGGCA	2728
QY	2809	TTGACGAGTGGTGGAGGAATGAACAGTTCTGGGTCAATTGAGGATPACTCTGCAACATCTGT	2868
Db	2729	TCGATGAATGGTGGAGGAATGAGCAGTTTGGGTATTGGAGGTATTTCCGCTCATCTAT	2788
QY	2869	TTGCGCTCTTTTCAGGGTCTTTCTGAAAGTGCTTCGCGGTATCGACACCACTTCACTGTCTCA	2928
Db	2789	TTGGGTCTTCCAGGGTCTGCTGAAGTGCTTGTCTGGAATTGATACCAGCTTCACTGTCA	2848
QY	2929	CCTCAAGGCTAATGAGAAAGGGCACTTTGCTGAGCTCTACATGTTCAAAGTGACGA	2988
Db	2849	CCTCAAGGGCTCTGATGAAGAAAGGGCAATTCGCTGAGCTCTACATGTTCAAAGTGACAA	2908
QY	2989	CGCTTCTCATCCCTCCGACGACCATTTTGATCATTAACATGGTTGGTGTGCTGGTGGCA	3048
Db	2909	CCCTCTTGATCCCAACGACCACTATTTTGATCATCAACCTTGGTCGGTGGTTGTTGCTGGTA	2968
QY	3049	CCTCTACGCCATCAACAGTGGTTTACCAATCATGGGGGGCGCTCTTTGGGAAGCTCTTCT	3108
Db	2969	TCTCTTAGCTATCAACAGTGGTACACGATCATGGGTCCGCTCTTTGGGAAGCTCTTCT	3028
QY	3109	TTGCTCTCTGGGTGATTGTTCACTTATPACCCATTCCTCAAGGGTCTTATGGGCGAGGAAA	3168
Db	3029	TTGCGTTCTGGGTGATTGTCCACCTGTACCCCTTCTCTTAAGGGTCTCATGGGTCCGCGAGA	3088
QY	3169	ACCGCACACGACGATTGTTCATCGTCTGGGCTGTCTCTCTGCTTCTATCTTCTCCTTGC	3228
Db	3089	ACCGTACACGACCATTTGTTGTTTGGGCCATCTCTCTTGTGATCGANTCTTCTCCTTGC	3148
QY	3229	TGTGGGTTTCGTGTGATCCATTCACTACCCGTCTCGCTGGCCCCAAATATCCAAACCTGTG	3288
Db	3149	TGTGGGTTCGATCGATCCATTACCAACCGGTGTCAAGGCCCTGTATCCAAAATGTG	3208
QY	3289	GCATCAACTGCTAG 3302	
Db	3209	GTATCAACTGCTAG 3222	
RESULT 9	AF200529	3676 bp	linear
LOCUS	AF200529		PLN 31-AUG-2000

RESULT 9
AF200529
LOCUS

Db 377 TTCCCGTGTGCGCCATGCTACGAGTACGAGCGCAAGCGGCGTGC 436
QY 246 CAGTGCAGACCAAGTACAAAGCCCAAGAGGAGCCAGCGGAGAGG 305
Db 437 CAGTGCAGACTAAGTACAAAGCCCAAGAGGAGCCAGTACACGCTGAGGAAAT 496
QY 306 GAGGACACTGATGCCATGATGTAGTACTCACTACCTGCACTGAGGAC 365
Db 497 GAGGATGTGATGCTGACGATGTAGTACTCACTACCAAGCATCTGCAACACCGGAT 556
QY 366 CAGAAGCAGAAGATGTGACAGGATGCGCAGCTGGCGATGAACACCGGGGCGAGTGGC 425
Db 557 CAGAGCAAAAGATGCTGAGAGATGCTCAGTTGGGAGCAAACTCAGTGGCAGTG-- 614
QY 426 AATGTTGGCCACCCCAAGTANGACAGTGGGAGATCGGCTCTCCAAAGTANGACAGTGA 485
Db 615 -ATATTGGGCTGGCTAAGTATGACAGCGGTGAATTTGGGCATGGGAAGTATGACAGTGGT 673
QY 486 GAGATCCCTAGGGATACGTCCTTCAGTCACCAACAGCCAGATGTGAGGAGAAATCCCT 545
Db 674 GAGATCCCTGCTGGATATATCCCGTCACTAATCTATAGCCAGATCTCAGAGAGATTCCT 733
QY 546 GAGGCTTCGCTCATCATCATGATGCTCCCTACGGGGAACATCAGCAGACGTGCTCCG 605
Db 734 GAGGCTTCCTCATATGATGCTCCTGTTGGGAACATTTGGCGAGGCGTGACATCAA 793
QY 606 TTTCCTATGTGAATCATTCACCAATCCGTCGAAGGAGTCTCCGGCAGTATTTGGGAAT 665
Db 794 TTTCCTATGTAAATCATTTCCAAACCCATCGAGGAGTCTCCCGTAGCCTTTGGCAAT 853
QY 666 GTTGCTCGAAACAGAGAGTTGATGCTGGAAATGAAGCAGCAGCAGGCTGGATTCCT 725
Db 854 GTTGCTCGAAACAGAGGTTGATGCTGGAAATGA--AGGATAAGGTGCATTCCT 910
QY 726 ATGACTAATGGGCAAGCAATGCTCCCTCTGAAGGTGGGAGCTACTGACATCGATGCA 785
Db 911 ATGACCAATGGAACAAGCAATGCTCCATCAGAAGGCGTGAGTGTGCTGATATGATGCT 970
QY 786 TCTACTGATACAACTAGGAGAGCTTTACTGAATGATGAATCGCCAGCTCTATCT 845
Db 971 TCTACTGATATACATGGAAGATGCTTACTGAATGATGAATCGGCAACCTCTATCT 1030
QY 846 AGAAAGTCCCATGCTCCCTCCAAATAATCCCTACAGATGCTGATGCTGCGG 905
Db 1031 AGAAAGTCCCATGCTCCCTCCAAATAATCCCTACAGATGCTGATGCTGCGG 1090
QY 906 TTGGTTGTTCTAAGCATCTTCCTGCACCTACCGTCTCACAAATCCTGTGGGTAATGCATAC 965
Db 1091 TTGGCTGTTCTATGCATATTTCTGGCTACCGTATCACACATCCTGTGAACAATGCATAT 1150
QY 966 CCACTGTGGCTTTATCTGTTATATGTGATTTGGTTGCTTTATCTCGGATCTGGAT 1025
Db 1151 CCACTGTGGCTTTATCCGTCATATGTGATCTGGTTGCTTTGCTGCTGGATTTGGAT 1210
QY 1026 CAGTCCCGAAGTGTTCCTCAATCAACCGGAGACCTACCTGATAGACTGGCTTTAAGG 1085
Db 1211 CAGTCCCGAAGTGTTCCTCAATCAACCGGAGACATACCTGATAGACTGGCTTTAAGG 1270
QY 1086 TATGACCGAAGGTGAACCGCTCAGTTGGCTGCTGTTGACATATTTGTGACAGCTC 1145
Db 1271 TATGACCGAAGGTGAACCGCTCAGTTAGCTCTCTGATATTTTGTGACAGTCTGTG 1330
QY 1146 GRACCTTTGAGAGCGCACCTATCTGCTACCTGCAACACTGCTGCTATCCATCTTCTGTT 1205
Db 1331 GATCCAAATGAGAGCGCTCTCTGCTACCTGCAAAATCTGCTGCTTCCATCTTCTGCTC 1390
QY 1206 GATTATCCCGTGCACAGGCTCTCTGCTATGATCTGATCAGCGAGCTTCAATGCTGACT 1265
Db 1391 GATTATCCCGTGCACAGGCTCTCTGCTATGTTTGGATGATGAGGCTGCTATGCTGACT 1450
QY 1266 TTTGACGCATTTGGCTGACATTCAGAGTTGCTAGGAAATGGGTACCATTTGTGAAGAAG 1325

Db 1451 TTTGATGCTCTCTCTGAAACTTTAGAGTTTGTAGAAAATGGGTTCCGTTCTCTAAGAAG 1510
QY 1326 TATGACATTTGAACCCAGAGCTCCGAGTTTACTTTTCCAGAAAATGATTAACCTGAAA 1385
Db 1511 TACAACATAGAGCTTANGCCCGGAAATGGTACTTTGCTCAGAAAATGATTAACCTGAAA 1570
QY 1386 GACAAAGTCCAGCTTCAITTTGTTAAAGACCCCGGCGCATGAAGAGAAATATGAAGAA 1445
Db 1571 GACAAAGTCCAGCTTCAITTTGTTAAAGACCCCGGCGCATGAAGAGAAATATGAAGAA 1630
QY 1446 TTTAAATCAGGATAAATCCCTAGTTTCTAAGGCATTTGAAAGTCCCGGAGGAGATGG 1505
Db 1631 TTTAAAGTTCGTTATCAATGCTCTTTAGCCAAAGGACAAAAGTTCCCGAGGAGGATGG 1690
QY 1506 ATCATGCAAGATGGCACACATGCGCCAGGAAACATACAGGATCATCTCTGGAATGAT 1565
Db 1691 ATCATGCAAGATGGTACACCTTTGGCTGGGAAACATACAGGATCATCTCTGGAATGAT 1750
QY 1566 CAGGTTTCTTGGTCCACAGTGGTGGCTTTGATCTAGAGGTAATGAGCTCCCGGCTTTA 1625
Db 1751 CAGGTTTCTTGGTCCACAGTGGGAGGCTTTGAGTTGAAGGCAATGAATCTCTCTGCTT 1810
QY 1626 GTTATGTGCTCGTGAAGAGCTCTCGGTTCCAGCACCAAGAGGCTGTGCGCATG 1685
Db 1811 GTTATGTGCTCGTGAAGAGCTCTCGGTTCCAGCACCAAGAGGCTGTGCGCATG 1870
QY 1686 AATGCCCTTGTGCTGCTCTCAGCTGCTTACTTAATGGAACAATACATGTTGAATCTTGAT 1745
Db 1871 AATGCCCTTGTGCTGCTGCTCTCAGCTGCTTACTTAATGGAACAATACATGTTGAATCTTGAT 1930
QY 1746 TGTGATCACTACATCAACCAACAGCAGGCTGTCCGAGAGCTATGTGCTTCTTAATGGAT 1805
Db 1931 TGTGACCATCATCATCAATATAGCAAGGCTCTTCGAGAGCTATGTGCTTCTTATGGAC 1990
QY 1806 CMAAAGTGGTCCGCAAGTCTGTGTTAGTGCAGTCTCCCAAGAGTTTCATGGGATGAT 1865
Db 1991 CMAAAGTGGTCCGCAAGTCTGTGTTAGTGCAGTCTCCCAAGAGTTTCATGGGATGAT 2050
QY 1866 AGGAATGATCGATATCAACCAACAGCAGGCTGTCCGAGAGCTATGTGCTTCTTAATGGGCTT 1925
Db 2051 AGGAATGATCGATATCAACCAACAGCAGGCTGTGTTAGTGCAGTCTCCCAAGAGTTTCATGGGCTT 2110
QY 1926 GAGCGCATTTCAAGGACAGCTTATGTGGAACTGTTGTTTTCACAGCAACAGCTATC 1985
Db 2111 GAGCGCATTTCAAGGACAGCTTATGTGGAACTGTTGTTTTCACAGCAACAGCTATC 2170
QY 1986 TATGTTATGAGCCCGCAATTAAGCGCAAGAGCAGGTTTCTTGGCATCACTATGTTGG 2045
Db 2171 TATGTTATGAGCCCTCCAGTCAAGAAAAGGAGGCTTCTTCTTGGCTTTGTTGGG 2230
QY 2046 GCAAGAGAGCAAGCAAGTCAAGAAAAGGAGCTCAGATAAGAAAAGTCCGAAAG 2105
Db 2231 GCAAGAGAGCAAGCTCAAAATCTAAGAA--GAGCTCGAAAAGAGAGTCACTAGAA 2287
QY 2106 CATGTGACAGTCTCTGTTCCAGTATTTCAATCTCGAAGACATAGAGGAGGCTGTTGAAGGT 2165
Db 2288 CAGCGACAGCTCTCTGACAGTATTTAATCTCGAAGATATAGAGGAGGATTTGAAGGT 2347
QY 2166 GCTGGGTTGATGATGAGAAATCAGTCTCATGTCTCAATGAGCTTAGAGAGAGATTT 2225
Db 2348 TCTCAGTTTGTATGAGAAATCGCTGATTTATGTCTCAATGAGCTTTGGAAGAGATTT 2407
QY 2226 GCGCAGTCAAGCAGCTTTGCTGCTCCACTCTGATGGAATATGCTGGTGTCTCTCAGTCC 2285
Db 2408 GCGCAGTCAAGTGTGTTGTTGAGCTCTCTCTGATGGAATATGCTGGTGTCTCCACATCT 2467
QY 2286 TCCACTCCAGAAATCTCTTTTGAAGAGAGCTATCCATGTCTAAGTTGTGGCTATGAGGAC 2345
Db 2468 GCAACTCCAGAGTCTCTTCTGAAAGAGAGCTATTTCAATGTCTATGCTGTGGCTATGAGGAC 2527
QY 2346 AAGTCTGATGGGAACTGAGATTTGGTGGATCTATGGATCTCTCAGAGAGATATCTT 2405
Db 2528 AAAAAGTGGGAACTGAGATTTGGTGGATCTATGGTGTCTGTTTACAGAGACATTTCTC 2587

QY 2406 ACTGATTCAGATGACGACGAGGCTGGCTTTCAGTCTATTGTCATGCCCAAGCGCCA 2465
|||
Db 2588 ACCGATTCAGATGATGATGCTGCGAGGCTGGGATCAATCTACTGTCATGCTTAAGGACCA 2647
|||
QY 2466 GCTTTCAAGGGATCTGCCGCCCAATCAATCTTTCAGATCGTGAACCAAGTGTGGGTGG 2525
|||
Db 2648 GCTTTCAAGGGATCTGCTCTCAATCAACCTTTCGATCGTTTGAATCAAGTGTGGGTGG 2707
|||
QY 2526 GCTCTGGTCTGTTGAAATCTTTTCAGCGGCAATGCGCCCTTATGATGATGCTTACGGA 2585
|||
Db 2708 GCTCTGGTCTCAATGAAATCTTTTCAGCAGCAATGTCCTCAATGATGATGCTTATGGA 2767
|||
QY 2596 GGGCGCTCAAGTCTCGAGAGATTCGCTTACATCAACACCAACATTTACCCATAC 2645
|||
Db 2768 GGGCGCTTAAATCTCGAGAGATTTGCTTATATCAACAAATTTATCCATCACA 2827
|||
QY 2646 TCTCTCCGCTTCTAGTCTATTGATATTGCTGCTATCTGCTCTCACTGGAAGTTC 2705
|||
Db 2828 TCAATCCGCTCTCTGCTACTGATATTGCGAGATTTGCTTCTCACTGGAGTTC 2887
|||
QY 2706 ATCATGCCAGATTAGCAACTTGGCCAGTATCTGGTCAATGCGCTCTTCTTCAAT 2765
|||
Db 2888 ATCATCCAAAGATTAGTAACTAGAGAGTGTGTTGTTTATATGCTCTTTATCTCAAT 2947
|||
QY 2766 TTCGCCACTGTTATCTTGAGATGAGTGGAGTGTGTTGCAATGAGAGTGTGGAGG 2825
|||
Db 2948 TTGCGCACTGGTATCTTGAGATGAGTGGAGTGTGTTGCAATGAGTGTGGAGG 3007
|||
QY 2826 AATGAACAGTTCTGGGTCAATTTGGAGTATCTGTCACATCTGTTTGGCTCTTTCAGG 2885
|||
Db 3008 AACGAGCAGTTCTGGGTCAATTTGGTGTATTTCTGCGCATTTATTTGCGCTCTTCAGG 3067
|||
QY 2886 CTTCTGAAGGTGCTTGGGTATGACACCAATTTCACTGTCACCTCAAGGCTTAATGAC 2945
|||
Db 3068 CTCCTGAAGGTGCTTGGGTATGACACAGAGCTTCACTGTCACCTCAAGGCTTAATGAC 3127
|||
QY 2946 GAAGAAGCGATTTGCTGAGCTTACATGTTCAAGTGGAGAGAGCTTCTATCCCTCG 3005
|||
Db 3128 GAAGAAGGTGATTTTGGCGAGCTTACAGTTTCAAGTGGAGAGAGCTTCTATCCCA 3187
|||
QY 3006 ACGACCAATTTTATCAATTAACATGTTGGTGTGCTGTCGCTCTTCAAGGCTTCTTCA 3065
|||
Db 3188 ACCACTATTTTATCAATCAACCTGCTGCGGTGCTGCTGCTGCTTCTTCAAGGCTTCT 3247
|||
QY 3066 AGTGTACCAATCATGCGGCGCTCTTTGGAGAGCTTCTTCTTGGCTCTTGGGTGAT 3125
|||
Db 3248 AGCGTTTACCAATCATGCGGCGCTCTTTTGGAGAGCTTCTTCTTGGGTGAT 3307
|||
QY 3126 GTTCACTTATACCAATCTTCAAGGCTTCTTATGGGAGAGAGAGAGAGAGAGAGAG 3185
|||
Db 3308 GTCCACTGTACCTTCTTCAAGGCTTCTTATGGGAGAGAGAGAGAGAGAGAGAGAG 3367
|||
QY 3186 GTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3245
|||
Db 3368 GTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3427
|||
QY 3246 CCATTCATACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3304
|||
Db 3428 CCATTCACACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3486
|||

RESULT 10
AY055724
LOCUS 3532 bp mRNA linear PLN 16-DEC-2003
DEFINITION Populus tremuloides cellulose synthase (Cesa5) mRNA, complete cds.
ACCESSION AY055724
VERSION AY055724.2 GI:39933009
KEYWORDS Populus tremuloides (quaking aspen)
SOURCE Populus tremuloides
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
FEATURES
source

gene

CDS

rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
1 (bases 1 to 3532)
Joshi,C.P. and Kalluri,U.C.
Isolation of two full-length Cesa genes from aspen trees homologous
to primary cell wall related Cesa genes from Arabidopsis
Unpublished
2 (bases 1 to 3532)
Joshi,C.P., Chavli,R., Kalluri,U.C. and Samuga,A.
Hypervariable II Regions from New Members of Aspen Cellulose
Synthase Superfamily Aid in Clarifying their Phylogenetic
Relationships with Other Cellulose Synthases in plants
Unpublished
3 (bases 1 to 3532)
Joshi,C.P. and Chavli,R.
Direct Submission
Submitted (11-SEP-2001) Forestry, Michigan Tech. University, 1400
Townsend Drive, Houghton, MI 49931, USA
4 (bases 1 to 3532)
Joshi,C.P. and Kalluri,U.C.
Direct Submission
Submitted (10-OCT-2002) Forestry, Michigan Tech. University, 1400
Townsend Drive, Houghton, MI 49931, USA
Sequence update by submitter
On Dec 16, 2003 this sequence version replaced gi:23428423.
Location/Qualifiers
1..3532
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
1..3532
/gene="Cesa5"
/note="PtrCesa5"
236..3472
/gene="Cesa5"
/codon_start=1
/product="cellulose synthase"
/protein_id="PAL23710.2"
/db_xref="GI:39933010"
/translation="MESEGETGAKPMKSTGGQVCOI CGDNVGTADGPFVACDVCAF
PVCPCYEYERKDNQSCPCQKTRYKLGSPAILGDEEDGADGASDFNYSENQ
QKQIABRMKSWMTYGPEDSGASNDYKEVSHNHI PLITNGHGVSGDLSAEPHIS
MASPGAGGKRI PVASDVHSSNVVDPVPEFGSPGLGNVAKERVDMGKQDKTV
VPMSTGHAPSERGAGDIDAADVLVDLSLNDARQPLSRKVSIPSSKINPYMVIVL
RLGILCLPLHRIITVFNAYALGILYIGWNGFALSKILLQDFPKNRETYDDRL
ALRYDMGEPHSLVVVDIFARSGVHLKEPPLVPANAVLSILAGSDPIDKVSVDG
AAMUTFALSETSEFSRWKVFCKYSIEPRAPWYFAQIDYLDKQVQPSFVKDRA
MKREYEFKIRINGLAKAQVPEGMIMQDTEWPGNNTDHPGMIQVFLGSGGLD
SDGNELPRLVVSREKRPFOHKKAGAMNSLVRSVAVLTNGPFLNLDGCHYINISK
ALREAMCFMDPNLGHVCYVQRFQDFGIDRNDRYANRTVFFDINLRGLDGTQGPV
YVGTGVFNRTALYGYEPPLKPKHKFQMLSSLCGSRKSKSGSKSGKSGKHV
DPTVPIFSLDDIEBEGVAGDFDDKSLMSQMSLEKRPQSAFVASTLMENGVPQS
ATPTELLKAHIVISCGYEDKTDWSEIGTWYGTVEDILTFGKMHARGMSIYCMXP
RPAFGKAPINLSDLNLKWLKALSGVEILLSRHCPWYGYGRLKWLRFAYVNTTI
YPTAIPALPYTTPALICLLDKFIIPQISNIAIWFISLFLSIPATGILLBRWSGVG
IDEMWNEQFWIIGVSAHLPAVFGLLKVLADITNTFTVTSKASDEGDSAEYLFX
WTLILPTEILLIIVNLGVVAGISHAINSQSWGPLEGKLFPAPWIVHLYFLPKGL
MGRQNRPTTIVVWMSILLASIFSLMLWRVDFITRTVTPDVEQCGINC"

ORIGIN

Query Match 47.4%; Score 1720; DB 8; Length 3532;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 2332; Conservative 0; Mismatches 875; Indels 27; Gaps 6;
QY 83 GAAGTCGGGAGGACCGGGCGGGGACGTGTGCGAGATCTGCGCCGACGCGCTGGGCAC 142
|||
Db 262 GAAGCCCATGAAAGTACAGGTGACAGTTCGCCAGATCTGTGTGATATGTTGCAA 321
|||
QY 143 CACGTTGACGCGGACGCTTTCACCGCTGCGAGCTGCGCGCTTCCCGGCTCGCGGCC 202
|||
Db 322 GACAGCGGATGTTGAAACCGTTTGTGTTGCGATGTTTGTGATTTCCAGTGTGAGGCC 381
|||
QY 203 CTGCTACGACGACGCGCAGGAGGACCCAGCGCTGCTCCAGTGCAGACCAAGTA 262
|||

Db 382 TTGCTATGAGTACGAGAGGAGGATGGCAATCAGTCTTGCCCTCAGTCGCAAAACAGATA 441
QY 263 CAAGCCCAACAGAGGAGCCAGCGATCCCGGGAGGAGGAGCACTGATGCCGA 322
Db 442 CAAGAGGCTCAAGAGGTAGCCCTGCCATCTTGGTGATAGAGAGAGGATGGTATCTGA 501
QY 323 TGATGTAGTAGTACTCAACTACCTGATCTGATCTGAGGACGAGAGCAGAGGATGTC 382
Db 502 TGATG--GTGCTAGTAGTATTTAATTTACTCAGAAATCAAAACAGAAACAGAGATGTC 558
QY 383 TGACAGATGCGCAGCTGGCGCATGAACACCGGGGCGAGTGGCAATGTTGGCCACCCCAA 442
Db 559 GGAACGCATCTGAGCTGGCAGATGACTTATGACACAGGGGAAGATTCTGGAGCTTCAA 618
QY 443 GTATGACAGTGGCGAGATGGCCCTCTCAAGTATGACAGTGGAGAGATCCCTAGGGGATA 502
Db 619 TTATGATAAAGAGG---TTTCGCATAAACCATATCCACTACTCACCATAATGGACATGAGG 674
QY 503 CGTCCCTTCAGTCACCAACAGCCAGATGTCAGGAGAAATCCCTGGAGCTTCGCTGATCA 562
Db 675 TTCTGGAGATCTGTCTGCTGATCACCTGAGCATATTTCTATGGCCTCTCCCTGGTGTG 734
QY 563 TCACATGATGTCCTTACCGGGAACATCAGCAGACGTGCTCGTTCCTTATGTGAATCA 622
Db 735 GTGGTGGGAAGCGATCCCTTATGCACTCTGATGTCATCAATCATCTAATGTGAGAGTTG 794
QY 623 TTCACCAATCCGTCAAGGAGTCTTCGCGCAGTATTTGGAAATGTTCCCTGGGAAAGAG 682
Db 795 TGGACCCAGTAAGGAGGTTTGGTTTACCGGGAGCTG--GGCAATGTGGCCCTGGGAAGAA 852
QY 683 AGTTGATGGCTGAAATGAAGCAGCAAGGAGGTGGATTCCTCATGACTAATGGGACAAG 742
Db 853 AGTTGATGGCTGAAAGATGAAGCAGCAAGACTGTTTCCCTATGAGTACTGG----- 906
QY 743 CATTTGCTCCCTCTGAAGGTGGGAGCTACTGATCGATCGATCTACTGAATPACAAAT 802
Db 907 ---TCATGCTCCTCTGAAAGGGGTGCTGGAGATATTTGATGCTACGGATGTGCTGT 963
QY 803 GGNAGAGCTTTTACTGAATGATGAAGTCCGAGCCCTCTATCTAGAAAGTCCCGATTGC 862
Db 964 GGATGACTCTTTACTGAATGATGAAGTCCGCAACCTCTTTCGAGGAAGTTTCCCATTC 1023
QY 863 TTCCCTCAAAATAAATCCCTACAGATGGTCAITTTGTTCTGCGGTGTTGTTCTTAAGCAT 922
Db 1024 TTCTTTAGATTAATCCATACAGGATGGTTATTTGTTTGGCGCTTGGTATCTCTGCA 1083
QY 923 CTTCCTGCACTACCGTCTCAAAATCTGTGCGTATGCAATCCCATGTGGCTTTTATC 982
Db 1084 TTTCTTTGCACTACCGAATAACAAATCTGTGCGTAAATGCCATGCTCTGGGGTTATATCT 1143
QY 983 TGTATATGTGAGATTTGTTGTTTATCTCTGGATCTGGATCAGTTCCCGAAGTGGTT 1042
Db 1144 GGGGATATGGGAGATTTGGTTTGAATATCCCGGATCTGGATCAGTTCCCTAAGTGGCT 1203
QY 1043 TCCAATCAACCGGGAGACCTACCTTTGATAGACTGGCTTTAAGGTATGACCGAGAGGTGA 1102
Db 1204 CCCTGGAAATCGGGAACATATCTTGACAGCTTGCTTTGAGATATGACATGGAAGAGA 1263
QY 1103 ACCGTCTCAGTGTGGCTGCTGTGACATATTTGTGCTAGTACAGTCCGATGGAAGAGCC 1162
Db 1264 ACCCTCCACTTGGTGTGGACATTTTGGCCAGGTCCGGGTCCCATTTAAAGAGGCC 1323
QY 1163 ACCTATGCTCACTGCCAAACAGTGTCTATCCATTTCTGCTGTTGATTTATCCCGTGAGCAA 1222
Db 1324 CCCTTTGGTGCCAGCAATGAGTGTCTATTTCTTGAGGAGTCTCCCAATTTGACAA 1383
QY 1223 GGTCTCTTGTATGTATCTGATACGGAGCTTCAATGCTGACTTTTGGAGCATTTGGCTGA 1282
Db 1384 AGTCTCATGCTATGCTCTGATGATGGTGTGCAATGTTGACATTTGAAGCTTTGTCTGA 1443
QY 1283 GACTTCAGATTTGCTAGGAAATGGGTACCAATTTGTGAAGAGATGATGACATTTGAACCCAG 1342
Db 1444 GACTTCAGAGTTTTTCCAGGAAATGGGTGCCCTTCTGCAAGAGTACAGCATTTGAGCCCTG 1503

QY 1343 AGCTCCCGAGTTTACTTTTCCAGAAAATTGATTACCTGAAAGACAAAGTCCAGCCTTC 1402
Db 1504 GGCTCCAGAAATGGTACTTTTGCACAAAGATTTGACTACTTGAGGATAAAGTGCAGCAATC 1563
QY 1403 ATTGTGTAAGAGCCGCGGGCCATGAAGAGAGAAATATGAAGAAATTTAAATACGAGATAA 1462
Db 1564 TTTTGTCAAAGACCGGTAGAGCTATGAAGAGAGAAATATGAAGAAATTTAAATTCGTATCAA 1623
QY 1463 TGGCCCTAGTTTCTAAGGCATTTGAAAGTCCCGAGGAGAGATGGATCATGCAAGATGGCAC 1522
Db 1624 TGGGCTTGTGTGCAAGCGCAAAAGGTGCTGGAAGAGGATGGATTTATGCAAGATGGGAC 1683
QY 1523 ACCATGGCCAGGAAACAATACCAGGGATCATCTCGGAATGATTCAGGTTTCTCGTCA 1582
Db 1684 CCATTCGGCTTGGAAATTAACACAGAGACCATCAGGAATGATTCAGGTTTCTTGGGCCA 1743
QY 1583 CAGTGTGGCCCTTTGATPACTGAGGGTAAATGAGCTCCCGCTTTAGTTTATGTCTCGTGA 1642
Db 1744 AAGTGGAGGGCTTGAATCTGATGTTAATGAGCTTCGCGGTTAGTTTATGTTCTCGTGA 1803
QY 1643 AAGCGTCTTGGTTCAGCACCAAGAGGCTGTGCCATGAATGCCCTTGTCTGTGT 1702
Db 1804 AAGCGTCTTGGCTTCAGCACCAAGAGGCTGTGCCATGAATTCACCTTGTCTGTGT 1863
QY 1703 CTCAGCTGTCTTACTAATGGCAATACATGTTGAATCTTGATTTGTGATCCTACATCAA 1762
Db 1864 ATCAGCGTGTCTAATTAATGCGGCTTCTTGTGAAATCTTGATTTGATCCTACATAAA 1923
QY 1763 CAACAGCAAGGCTGTCGAGAGAGCTATGTCTTCTTAATGGATCCAAACCTAGTCCGCA 1822
Db 1924 CATTTAGCAAGGCTTTGAGAGAGCTATGTGTTTTCATGAAGGATCCCTAACCTTGGGAAACA 1983
QY 1823 AGTCTGTTATGTCAGTTCACCAAGAGTTTGAATGGATTTGATAGGAATGATCGATATGC 1982
Db 1984 TGTCTGCTATGTTTCACTTTCCAAAGATTTGATGCAATGACAGGAATGATCGATATGC 2043
QY 1883 AAACAGCAACACTGTCTTTTTCATATTAACCTTGAGGGCTTTGACGGCAATCAAGGACC 1942
Db 2044 CAATCGCAATACATGTTATCTTTGATATAAACTTGAGGGGTCTGGATGTTATCAAGGACC 2103
QY 1943 AGTTTATGTTGGAACTGGTGTGTTTCAACAGAAACAGCTATCTATGTTTATGAGCCCC 2002
Db 2104 TGTCTATGTGGTACTGATGTCTTTAAACAGAGCGCATGTTATGTTGTTATGAACCTCC 2163
QY 2003 -----AATTAGGCCAAGAGCGAGTTTCTTGGCATCACATATGTTGGGG---CAAGAA 2053
Db 2164 CCTCAACCCCAAGCATTAAGAAACAGGCAATGCTCTTCACTCTGTGGTGGATCAAGAAA 2223
QY 2054 GAAGGCAAGCAAGTCAAGAAAGAGAGCTCAGATAAGAAAAAGTCCGAAACAAGCATGTGGA 2113
Db 2224 GAAGGGTTCTAATCAAGCAAAAGGGAATCTGACAGAGAAATCTGGAAGCATGTAGA 2283
QY 2114 CAGTTCTGTTCCAGTATTCAAATCTCAAGACATAGAGGAGGGTGTGAAGGTGCTGGGTT 2173
Db 2284 TCCCACTGTACCCATTTTTCAGCTTGGACGATATAGAAAGAGGCTTTGAAGGAGCTGGATT 2343
QY 2174 TGAATGATGAAGAAATCAGTTCTCATGCTCAAAATGAGCTTAGAGAAGAGATTTGCGCAGTC 2233
Db 2344 TGAATGAGAGAGTCACTTCTCATGCTCAAAATGAGCTTGGAGAGAGGTTTGGCAGTC 2403
QY 2234 AGCAGCATTTGTTGCTCCACTCTGATGGAATATGTTGTTGTTCTCAGTCTCCACTCC 2293
Db 2404 AGCTGTTTTTGTGCTTCCACACTTATGGAGATGTTGGTGTCTTCTCAGTCTGCAACCC 2463
QY 2294 AGAATCTCTTTTGAAGAGCTATCCATGTCATTAAGTTGTGGCTATGAGGACAAAGTCTGA 2353
Db 2464 AGAACTCTTCTTAAAGAGCTATTCATGTCATTTAGTTGTGGGTATGAGGACAAACAGA 2523
QY 2354 ATCGGGAATGAGATTTGGTGTGATCTATGATCTGTACAGAAGATATTTCTTACTGGATT 2413
Db 2524 TTGGGATCTGAGATTTGGATGGAATCTATGTTCTCTTTACAGAAGATATTTCTTACTGGTT 2583

[illegible]

	Query Match	46.3%	Score 1679.4	DB 8	Length 3229
	Best Local Similarity	72.0%	Pred. No. 0		
	Matches 2333	Conservative 0	Mismatches 836	Indels 72	Gaps 8
QY	77	CGCCCTGAAGTCGGGAGGACGGGGCGGGACGTGTGCCAGATCTGCCGCGACGGCCT	136		
Db	21	CGCGGGAAGCCGATCAAGAAATTTCCGAGACATTGCCAGATCTGTAGTGACAAATG	80		
QY	137	GGGACACCGTTGGACGGCGAGCTCTTACCGCCTCGGACGCTGCGGCTTCCCGCTG	196		

QY	2345	CAAGTCGAATGGGAACCTGAGATTGGTTGGATCTATGATCTGTACAGAGATATTTCT	2404
Db	2241	TAAGTCAGATTGGGAATGGAGATTGGATGGATCTATGGTTCTGTGACAGAGATATTTCT	2300
QY	2405	TACTGGATTCAAGATGCACGCAAGAGCGCTGGGGTTTCAGTCTATTGTCATGCCCAAGGCC	2464
Db	2301	GACTGGTTTCAAAATGATGCCCGTGGATGGGATCCATTACTGATGCCCTAAGTTCC	2360
QY	2465	AGCTTTCAGGAGATCGGCCCATCATCATCTTTCAGATCGTCTGAACCAAGTCTCGGGTG	2524
Db	2361	AGCTTTCAGGAGTCTGCTCCCTATCAATCTTTCAGATCGTCTGAACCAAGTCTCGGGTG	2420
QY	2525	GGCTCTCGGTTCTGTTGAAATCTTTTCAGCGGATTCGCCCTATGTTGATGGTACGG	2584
Db	2421	GGCTTTAGTTCAGTTGAGATCTCTTCAGTCGGCAATTCCTATATGATGATGGTTACAA	2480
QY	2585	AGGGCGCTCAAGTTCCTGGAGAGATTCCGTTTACATCAACACACCAATTCACCACTAAC	2644
Db	2481	TGGGAGGCTAAATTTCTTGAGAGGTTTGGGTATGTGAACACCACTACCTATCAC	2540
QY	2645	CTCTCTCCGCTTCTAGTCTATTGTATATGCTGCTATCTGTCTGCTCACTGGAAAGTT	2704
Db	2541	CTCCATCTCTCTCATGTATTGTATGATGTCAGTGGCGGTTTCTCTTCCACCAACGAGTT	2600
QY	2705	CATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTTCATTGCGCTCTTCTTTCAT	2764
Db	2601	TATATATCTCTAGATTAGTAACATTTGCAAGTATATGGTTTCTGTCTCTTCTCTCCAT	2660
QY	2765	TTTCCGCACTGATTCCTTGAGATGAGTGGAGTGGTGTGTCATTGACGATGGTGAG	2824
Db	2661	TTTCCGCACTGATTAATAGAAATGAGTGGAGTGGGTAGGCATAGACGATGGTGAG	2720
QY	2825	GAATGAACAGTCTCGGTCATTGGAGGTATCTCTGCAATCTGTTTGGCGTCTTTCAGG	2884
Db	2721	AAACGAGCAGTTTGGGTCATTGGTGGAGTATCCGCTCATTTATTCGCTGTGTTTCAAG	2780
QY	2885	TCTTCTGAAGTGCTTGGCGGTATGCACCAACTTCACTGTCACTCACTCAAGGCTAAATGA	2944
Db	2781	TATCTCTAAGTCTTGGCGGTATTGACAAACTTTCAGATTAATCAAAAGCTTCAGA	2840
QY	2945	CGAAGAGCGCACTTGTCTGAGCTCTACATGTTCAAGTGGACGAGCTTCTCATCCCTCC	3004
Db	2841	TGAAGACGAGACTTTGCTGAGCTCTACTGTTTCAATGGACAACTTCTGATTCGCC	2900
QY	3005	GACGACCAATTTGATCATTAATGATGGTGTGCTGCTGGACCTCTCAGCCATCAA	3064
Db	2901	AACGACGCTGCTCATTTGTAAACTTAGTGGGAGTTGTTGCAGGAGTCTCTTATGCTATCAA	2960
QY	3065	CAGTGGTTTACCAATCATGGGGCGGCTCTTTGGGAAGCTCTTCTTGGCTCTCGGTCAT	3124
Db	2961	CAGTGGATACCAATCAUGGAGCACTCTTTGGTAAAGTTGTTCTTGGCTCTCGGTCAT	3020
QY	3125	TGTTCACTTATACCAATTCCTCAAGGGTCTTATGGGAGGCAAAACCGCACACCGACAT	3184
Db	3021	TGTTCACTTATACCAATTCCTCAAGGGTCTTATGGGTCGACAGAACCGGACTCTTACAT	3080
QY	3185	TGTCATCTCTGGGCTGCTCTCTGCTCTTCTTCTTCTCTTCTGCTGCTGCTGCTGCTGA	3244
Db	3081	TGTTGCTGCTGCTGCTGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	3140
QY	3245	TCCATTCTACTACCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3304
Db	3141	TCCCTTCTACTACCGGATCACTGGCCCGGACATTTCTGGAATGTGGAATCAACTGTTGAGA	3200
QY	3305	A 3305	
Db	3201	A 3201	

RESULT 13
AF027174
LOCUS

DEFINITION

Arabidopsis thaliana cellulose synthase catalytic subunit (Ath-B) mRNA, complete cds.

ACCESSION
AF027174

VERSION
AF027174.1

KEYWORDS
GI:2827142

SOURCE
Arabidopsis thaliana (thale cress)

ORGANISM
Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE
1 (bases 1 to 3682)

Arioli, T., Peng, L., Betzner, A.S., Burn, J., Wittke, W., Herth, W., Camilleri, C., Hofte, H., Plazinski, J., Birch, R., Cork, A., Glover, J., Redmond, J., and Williamson, R.E.

Molecular analysis of cellulose biosynthesis in Arabidopsis

Science 279 (5351), 717-720 (1998)

98111412

9445479

2 (bases 1 to 3682)

REFERENCE
Arioli, T.

Direct Submission

Submitted (29-SEP-1997) Plant Science Centre, Australian National

University, Acton, Canberra, ACT 200, Australia

Location/Qualifiers

FEATURES

source

1..3682

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/cultivar="Columbia"

/db_xref="taxon:3702"

1..3682

/gene="Ath-B"

256..3453

/gene="Ath-B"

/note="RSMI-like"

/codon_start=1

/product="cellulose synthase catalytic subunit"

/protein_id="AAC39336.1"

/db_xref="GI:2827143"

/translation="MESSEGTAGPKMKNIVPQTQICSDNVGKTVGDRFVACDICSF
PVCPCYVERKQDNQSPQCKTQKLGKSPATPGDKEDGLDGVGVENYVPOKEK
ISERMLGHLTRGKEEMGEQYDKEVSHNHLPLTSPQDTSGESAASPERLSVST
IAGKRLPKSYSDNVQSNRRIVDPVGLGNVAKWERVDGKMKQKNTGFTVQAASER
CGVDIADTDLADEALNDKARQPLPSRRINPFMVMILRLALYLCLELHYR
ITNPFAFALWLSVCEIWFALSWILDQPKWFFNVRETYLDRLALYDREGESQ
LAAVDIFVSTDPLKEPLVTANTVLSILAVDYPVKVSCYVDFDGAAMLSFEISAE
SEFARKVFCCKYSIEPRAPETFAAKIDYLDKDKVTSFVKDRAMKREVEEPKIRI
NALYSKALCKPEGVMQDGPWPNNTGDHPGMIOVFLGONGGLDARGNELPLRVV
SREKRPQHHKAGAMNALVRVSAVLTNGPFIILNCDHVINNSKALREAMCFMLDP
NLGQVCYVQPPQFDGIDKNDRYANRNTVFEDINLGLDSIQGVVYVGTGCVFNRTA
LYGEPPIKVKHKKPSLLSKCGSRKKSKAKKESDKKSGRHTDSTVPVFNLDIE
EGVEGAFDDEKALLMSLEKRFQGSVAVSTLMENGVPSPSATPENLKEAIHV
ISCYEDKSDMGMEIGHIYGVSTEDILTGFMAHGWRSIYCMFKLPFAFGSAPINLS
DRLNQRLWALGSVEILLFSRHCPITWYNGRLKFLERFAYVNTTTPITSTPLMYCT
LLAVCLTNQFIIPQISNIASIFLSLFLSTFATGILEMBSGVGIDENWNEQFWVI
GYSUHLFAVYFQGLKVLGADTNFTVTSKASDEGDFAEILYLFNTLLIPPTLLI
VNLGVVAGVSIAINSQVSNQPLGKLFVAFVWVIVHLYPLKGLMGRQNRTPITVVV
WSVLLASIFSLMWIRIDPFTSRVTGPDILEGGINC"

ORIGIN

Query Match	45.2%	Score 1675.2;	DB 8;	Length 3682;
Best Local Similarity	71.4%	Pred. No. 0;		
Matches 2348;	Conservative	0;	Mismatches 868;	Indels 72; Gaps 8;
QY	30	TCGGTGGGGTGGATCGCTCTGCCGCGCATGACGCGGACGCGGACGCGCTCAAGTCC	89	
Db	229	TCGGTGGAGCACTAAGTGACAATCGAAGGGGAAACCGCGGGAAGCCG	288	
QY	90	GGGAGGCAAGGGCCGGGACGTGTGCCAGATTCGCCCGACGCGCTGGGACCAACGTTG	149	
Db	289	ATGAAGAACAATTGTTCCGCAGACTTGCAGATCTGTAGTACAAATGTTGCAAGACTGTT	348	
QY	150	GACGGCAGCTTTCACCGCCTCGGACGTCTCGCGCTCTCCGGTCTCGCGCCCTGCTAC	209	

349 GATGGAGATCGTTTGTGGCTTGTGATATTTGTTCAATCCCAAGTTTGTGGCGCTTGCTAC 408 Db
210 GAGCAGCGGAGAGAGGACCCAGGGCTGCTCCAGTGCAGACCAAGTACAGCGC 269 Qy
409 GAGTATGAGAGGAAAGATGGGAATCAATCTGTCTCAGTGCAGAAACCAAGATACAGAGG 468 Db
270 CACAGAGGAGCCAGCGATCCGCGGAGGAGGAGCGACACACTGATCCGATGATGGT 329 Qy
469 CTCAGAGGTAGTCTCTCTATCTCTGATTAAGACGAGATGGCTTAGTGTGATGAGGT 528 Db
330 A---GTGACTTCAACTACCTGCACTGCGCACTGAGGACCAAGAGAGAGATTTGCTGAC 386 Qy
529 ACTGTTGAGTTCACTACCT-----CAGAAAGAGAAATTTTCAGAG 570 Db
387 AGATCGGACGTGGCGCATGAACACCGGGGAGTGGCAATGTTGGCCACCCCAAGTAT 446 Qy
571 CGGATGCTGGTTGGCATCTTACTCTGTGGGAAGGAGAGAAATGGGGGAACCCCAAGTAT 630 Db
447 GACAGTGGGAGATCGGCTCTCCAAATATGACAGTGGAGAGATCCCTAGGGGATACGTC 506 Qy
631 GATAAAGAG-----TCTCTCAATCATCTTCTCTCTCAGGACGACAAAG-- 679 Db
507 CTTTCACTCACCACAGCCAGATGTTCAGGAGAAATCCCTGGAGCTTGGCTGGAATCATAC 566 Qy
680 -----ATACTTCAGGAGATTTTCTGCTGCCTCACCTCAACGCTC 720 Db
567 ATGATGTCCCTACGGGGAACATCAGCAGACGTGCTCCGTTTCCT---ATGTGATCAT 623 Qy
721 TCTGTATCTTCTACTATCGCTGGGGAAAGCGCTTCCCTATTTCATCAGATGCAATCAA 780 Db
624 TCACCAATCCGTCAAGGAGTTCCTCGGAGTATTGGGAATGTTGCTCGAAAGAGAGA 683 Qy
781 TCACCAATTAAGAGGATTGTGGATCTGTGAGTCCGGAATGTAGCTTGAAGGAGAGA 840 Db
684 GTTGATGGCTGGAATTAAGACAGGACAAGGGTGGATCCCATGATTAATGGGACAAGC 743 Qy
841 GTTGATGGCTGGAATTAAGACAGGACAAGGATTAATGCTGCT-----TGTGAGCAGC 891 Db
744 ATTGCTCCCTCTGAAGTCCGGAGCTACTGACATCATGATCATCTACTGATACAAATG 803 Qy
892 CAGGCTCTTCTGAAGAGGTGGAG---TAGATATTGATGCGCAGCAGATATCCTAGCA 948 Db
804 GAAGACCTTTACTGATGATGAACCTCGCCAGCTCTATCTAGAAAGTCCCAATGCT 863 Qy
949 GATGAGCTCTGATGATGACGAGCAGGAGCTCTGTCAAGGAAGTTCATTTCT 1008 Db
864 TCCTCCAAATAAATCCCTACAGAAATGGTCAATGTTCTGGGGTGGTGTCTAAGCATC 923 Qy
1009 TCATCAGGATCAATCTCTACAGAAATGGTATTATGCTGGCGCTTGTATCTCTTGTCTC 1068 Db
924 TTCCTGCACTACGCTCTCAAAATCCCTGCGTAAATGATCATACCACTGTGGCTTTTATCT 983 Qy
1069 TTCTTGATTTACCGTATAACAAACCCAGTGCAGAAATGCTTGTCTATGCTGCTGTCT 1128 Db
984 GTTATATGTGAGATTGGTTTGTCTTTATCTCTGGATCTGATCAGTTCCTCCAGTGGTTT 1043 Qy
1129 GTGATATGTGAGATCTGGTTGCTTATCTCTGATTTTGGATCAGTTTCCCAAGTGGTTT 1188 Db
1044 CCAATCAACCGGGAGACTTACTGTGATAGATCGGCTTTAAGTATGACCGAGAGAGTGAA 1103 Qy
1189 CCTGTGAACCGTGAACCTACTCGACAGGCTTCTTTAAGATATGATCTGTGAAGGTGAG 1248 Db
1104 CCGTCTCAGTGGCTGCTGTTGATATTTCTGATCAGTCCAGCCCTTCAAGAGGACA 1163 Qy
1249 CCATCAGATTAGTCTGTTGATATTTCTGATGATCTGTTGATCCCTTGAAGAGGACA 1308 Db
1164 CCTATCGTCACTGCAACACTGTGCTATCCATTTCTGCTGTGATTTATCCGTGACAAAG 1223 Qy
1309 CCCCTTGTGACAGCCACACAGTGTCTCTATTTCTGCTGTGATACCCAGTTGACAAAG 1368 Db
1224 GTCTCTGCTATGATCTGATGACGAGCTTCAATGCTGACTTTGACGATTTGGCTGAG 1283 Qy
1369 GTGTCCTGTTATGTTTTGATGATGGTGTCTCTGTTATCATTTGAATCATCTTGCAGAA 1428 Db

1284 ACTTCAGAGTTTGTCTAGGAAATGGGTACCAATTTGTGAAGAGTATGACATTTGAACCAGA 1343 Qy
1429 ACATCAGAGTTTGTCTGTAATGGGTACCAATTTGCAAGAAATATAGCATAGAGCTCGT 1488 Db
1344 GTCCCGAGTTTACTTTTGGCAGAAATTTGATTAATCTCTGAAAGACAAAGTCCAGCTTCA 1403 Qy
1489 GCACCAGATGCTACTTTGCTCGAAATAGATTAATCTTGAAGGATTAAGTTTCAGACATCA 1548 Db
1404 TTTGTTAAAGACCGCGGCCATGAGACAGATATGAAGAAATTTAAATTCAGATAAAT 1463 Qy
1549 TTTGTTCAAGATGCTAGAGCTATGAAGAGGAAATGAGAAATTTAAATTCGAATCAAT 1508 Db
1464 GCCTAGTTTCTAAAGCATTTGAAAGTCCCGAGGAAAGGATGGAATTCATGCAAGATGCAACA 1523 Qy
1609 GCATCTGTTTCCAAAGCCCTAAATGCTCTGAAGAGGGTGGGTATGCAAGATGCAACA 1668 Db
1524 CCATGCGCAGGAAACAAATACAGGAGTATCTCTGGAATGATTCAGGTTTTCTTGTCTCAC 1583 Qy
1669 CCGTGGCTGGAATTAATACAGGGGACCATCCAGGAATGATCCAGGTCTTCTTAGGCA 1728 Db
1584 AGTGTGGCTTGATACTCAGGGTAAATGAGTCCCGCTTTAGTTTATGTGTCTCTGTAA 1643 Qy
1729 AATGGTGGACTTGTATGAGAGGCAATGAGTCCCGGTTGGTATATGTTTCTCGAGAA 1788 Db
1644 AAGCGTCTCTGGGTTCAGCACCACAAAGAGGCTGGTCCCATGAAATGCCCTTGTGTGTC 1703 Qy
1789 AAGCGACCAAGATTCAGCACCACAAAGAGGCTGGTCTATGAATGCACTGTGTGAGATT 1848 Db
1704 TCAGTGTCTTACTTAATGGAACAATCATGTTGAATCTGATTTGATGATCACTACATCAAC 1763 Qy
1849 TCAGCAGTTCTTACCAATGAGCCTTTCACTCTGAATCTTGAATGATCACTATCAAT 1908 Db
1764 AACAGAAAGCTCTCGAGAAAGTATGCTCTCTAATGATCCAAACCTAGTGTCCGCAA 1823 Qy
1909 AACAGAAAGCTTAAAGAGCAATGTCTTCTGATGACCCCAACCTCGGGAAGCAA 1968 Db
1824 GTCTGTTATGTGAGTTTCCCAAAAGGTTTGAATGGGATTTGATGGAATGATGCAATATGCA 1883 Qy
1969 GTTGTGTTATGTTCTCAGTTCCTCCCAAGATTTGATGGTATCGATAAGAACGATAGATGCT 2028 Db
1884 AACAGAAAGCTCTTTTGTATTAATCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCT 1943 Qy
2029 AATCGTAATACCGTGTCTTTGATATTAATCTGAGAGGTTTAGATGGATTCAGAGACT 2088 Db
1944 GTTATGTGGAACTGCTGTGTTTCAACAGAAACAGCTATCTATGTTGTTATGAGCCCCCA 2003 Qy
2089 GTATGTCGAACTGATGTTTCAACAGAAACAGCTATTAACGCTTATGAACTTCA 2148 Db
2004 ATTAAGG-----CGAAGAACCCAGGTTTCTTGGCATCACTATGTTGGGGCAAGAAAG 2057 Qy
2149 ATAAAGTAAACACAAAGAACCAAGTCTTTTATCTAAGCTCTGTGTGGATCAAGAAAG 2208 Db
2058 GCAGCAAGTCAAGAAAGAGGCTCAGATAAGAAAGTCCGAACAGCATGTGACAGT 2117 Qy
2209 AAGAAATTCAGAAAGTAAAGAGAGTCCGACAAAGAAATACAGCAGCATACTGACTCA 2268 Db
2118 TCTGTTCCAGTATTAATCTCGAAGACATAGAGAGGCTGTTGAAGGCTGGGTTGAT 2177 Qy
2269 ACTGTTCTGTATCACTCGATGATAGAGAGGAGTTGAAGGCTGGTGGTTGAT 2328 Db
2178 GATGAGAAATCAGTCTCATGTTCAAAATGAGCTTGAAGAGAGATTTGGCAGTCA 2237 Qy
2329 GATGAAAGGCGCTCTTAATGTCCAAATGAGCCTGGAGAGCGATTTGGAACGTCTGCT 2388 Db
2238 GCATTTGCTCCTCACTCTGATGGAATATGTTGTTGTTCTCTAGTCTCTCACTCCAGAA 2297 Qy
2389 GTTTTTGTGTTCTTACCCCTAATGAAATGCTGTTGTTCTCTCCTCAGCACTCCAGAA 2448 Db
2298 TCTCTTTGAAAGAGCTATCCATGTCTAATGTTGTGCTATGAGGACAAAGTCTGAATGG 2357 Qy
2449 AACCTTCTCAAGAGGCTATCCATGTTCATTTAGTTTGTGTTATGAGGATAAGTCAATGG 2508 Db

QY	2358	GGAACTGAGATTGGTTGGATCTATGATCTGTACAGAGATATCTTACTGTGAATCAAG	2417
Db	2509	GGAACTGAGATTGGATGATCTATGTTCTGTGACAGAGATATCTGACTGGTTCAA	2568
QY	2418	ATGACGCAAGAGGCTGGCTTCAGTCTATGTGATGCCAAGCCGACGCTTTCAGGGA	2477
Db	2569	ATGCATGCCGCTGGATGGGATCCATTACTGTCATGCCCTAAGCTTCAGCTTCAAGGCT	2628
QY	2478	TCGCCCCCATCAATCTTTTCAGATCTCTGACCAAGTCTCGGTGGGCTCTCGTTCT	2537
Db	2629	TCGCTCCCTCAATCTTTTCAGATCTCTGACCAAGTCTGAGTGGGCTTTAGTTCA	2688
QY	2538	GTGAAATCTTTTCAGCCGCAATCCCTCTATGTTATGGCTACGAGGGCGCCCTCAAG	2597
Db	2689	GTGAGATTCTCTCAGTCGCAATGCTCTATATGTTATGTTTACAAATGGAGGCTAA	2748
QY	2598	TTCCCTGGAGAGATTGCTTACATCAACACACATTTTACCCATAACCTCTCTCCGCTT	2657
Db	2749	TTTCTTGAGAGTTTTCGCTATGTGAACACCACTATACCTTATCCTCCATCTCTCT	2808
QY	2658	CTAGTCTATTGTATATTGCTGCTATCTGCTCACTGGAAGTTTCATGTCAGAG	2717
Db	2809	CTCATGTATTGATGTTAGCGTTTGTCTCTTCCACCAACAGTTTATTTCTCTAG	2868
QY	2718	ATTAGCACTTGGCCAGTATCTGGTTTCATTGGGCTCTTCTTTCAATTTTCGCCACTGT	2777
Db	2869	ATTAGTAACTTGAAGTATATGTTTCTGCTCTCTCTTCTCTCCATTTTCGCCGCTT	2928
QY	2778	ATCCTTGAGATGAGTGGAGTGGTTGGCATTTGAGAGTGTGGAGGAATGAACAGTTTC	2837
Db	2929	ATACTAGAAATGAGGTGGAGTGGCTAGCATAGACGAATGTTGGAGAAACGAGCAGTTT	2988
QY	2838	TGGGTCATTGGAGGATCTCTGCATCTGTTGCGTCTTTCAGGGTCTTCTGAGGTG	2897
Db	2989	TGGGTCATTGGTGGAGTATCCGCTCAATTTATTTGGCTGTTTCAAGGTTATCTCAAGTC	3048
QY	2898	CTTGGCGGTATCGACCAACTTCACTGTCACTCAAAGGCTAATGACGAAGAGCGAC	2957
Db	3049	CTTGGCGGTATTGACACAACTTTCAGTTACTCAAAGCTTTCAGATGAACGCGAGAC	3108
QY	2958	TTTGTCTGAGCTTACATGTTTCAAGTGGAGCGTCTCTCATCCCTCCGACGCCATTTTG	3017
Db	3109	TTTGTCTGAGCTTACTTGTTCAAATGGACCACTTCTGATTCGCGCAACGAGCTGCTC	3168
QY	3018	ATCATTTACATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3077
Db	3169	ATTGTAACTTATGTTGGAGTGTGTCAGGAGTCTCTTATGCTATCAACAGTGGATACCA	3228
QY	3078	TCATGGGGCCGCTCTTTGGGAAGCTCTTCTTTGCTTCTGCTGCTGCTGCTGCTGCTG	3137
Db	3229	TCATGGGGACCACTCTTTGGTAAATGTTCTTTGCTTCTGCTGCTGCTGCTGCTGCTG	3288
QY	3138	CAATTCCTCAAGGCTTATGGGAGGCAAAACCGACACCGACGATGTCATGCTGCTG	3197
Db	3289	CTTTTCTCAAGGCTTATGGGAGGCAAAACCGACACCGACGATGTCATGCTGCTG	3348
QY	3198	GCTGCTCTCTGCTTCTATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3257
Db	3349	TCTGTTCTCTGCTTCTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3408
QY	3258	CGTCTGCTGCGCCCAATATCAAACTCTGTGCTATCACTGCTAGGAA	3305
Db	3409	CGAGTCACTGGCCCGGACATCTGGAATCTGGAATCACTGTTTGAA	3456
RESULT 14			
AR267559			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			

ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 3614)
TITLE	Arioli, A., Williamson, R.E., Betzner, A.S. and Peng, L.
JOURNAL	Manipulation of cellulose and/or .beta.-1,4-Glucan
FEATURES	Patent: US 6495740-A 9 17-DEC-2002;
source	Location/Qualifiers
	1..3614
	/organism="unknown"
	/mol_type="genomic DNA"
ORIGIN	
	Query Match 46.2%; Score 1673.6; DB 6; Length 3614;
	Best Local Similarity 71.4%; Pred. No. 0;
	Matches 2347; Conservative 0; Mismatches 869; Indels 72; Gaps 8;
QY	30 TCGGTGCGGTTGGATCGCTCTGCGCGCCATAGTGCAGCGGACGCGGACGCGCTGAAGTCC 89
Db	190 TCGGTGTTGGAAGCAACTAAGTGACAAATGGAATCCGAGGAGAAACCGCGGAAAGCCG 249
QY	90 GGGAGGACGCGGCGCGGAGACGTGTGCGAGATCTGCGCGGACGCGCTGGGACACAGTTG 149
Db	250 ATGAAGAACATTTGTTCCGAGACTTGCAGACTCTGTAGTGACAAATGTTGGCAAGACTGTT 309
QY	150 GACGCGGACGCTTTCACGCGCTGCGAGCTCTGCGGTTCCCGGTTCTGCGGCTTCCGCTCTAC 209
Db	310 GATGAGATCGTTTGTGGCTTGTGATTTTGTTCATTTCCCGAGTTTGTGCGGCTTGTCTAC 369
QY	210 GAGCAGGCGGCAAGGAGGCGCACCCAGGCTCCCTCCAGTGCAGAACCAAGTACAGCGC 269
Db	370 GAGTATGAGAGGAAAGATGGGAATCAATCTTGTCTCAGTGCAAAACNAGATACAGAGG 429
QY	270 CACAGAGGAGCCAGCGATCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 329
Db	430 CTCAAAGGTAGTCTGCTGCTATTTCTGCTGATATAAGACGAGGATGGCTTAGCTGATGAGGT 489
QY	330 A---GTGACTTCAACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 386
Db	490 ACTGTTGAGTTCAACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 531
QY	387 AGGATGCGCAGCTGGCGCATGAACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTAT 446
Db	532 CGGATGCTTGTGTCATCTTACTGCTGCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 591
QY	447 GACAGTGGCGAGATCGGCTCTTCCAAATGATGACAGTGGAGAGATCCCTAGGGGATACGTC 506
Db	592 GATTAAGAGG-----TCTCTCACAATCATCTTCTCTGCTCTCACGAGCAGACAAG-- 640
QY	507 CCTTCAGTCAACACGCGAGATGTCAGAGAAATCCCTGGAGCTTCGCTGATCATCAC 566
Db	641 -----ATACCTTCAGGAGATTTTCTGCTGCTCTACCTTGAAGCGCTC 681
QY	567 ATGATGCTCCCTTACGGGAACATCAGCAGAGCTGCTGCTGCTTTCCTC---ATGCTGAATCAT 623
Db	682 TCTGTATCTTCTACTATCTGCTGGGGAAGCGCTTCCCTATTCAATGATGTCATCA 741
QY	624 TCACCAATTCGCTCAAGGAGTCTTCCGCGAGTATTTGGGAATGTTGCTTGGAAAGAGAGA 683
Db	742 TCACCAATATGAAGGATTTGGATCCTCTTGGACTCGGGAATGTAGCTTGAAGAGGAGAGA 801
QY	684 GTTCATGCTCGGAAATGAACGAGGACCAAGGGTGGGATTTCCCATGACTAATGGGACAAGC 743
Db	802 GTTCATGCTCGGAAATGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 852
QY	744 ATTGCTCCCTCTGAAGGTCGGGCGAGTACTGACATCGATGCTACTGTAATACACATG 803
Db	853 CAGGCTGCTTCTGAAGAGGTGGAG---TAGATATTGATGCGCAGCAGATATCTTAGCA 909
QY	804 GAAGACGCTTTTACTGAATGATGAACCTCCGAGGCTCTATCTAGAAAAGTCCCCATTGCT 863
Db	910 GATGAGGCTCTGCTGAATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 969
QY	969 TCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1023

Db 970 TCAACGGATCAATCTTACAGAAAGTATTAATGCTGGGCTGTGTTATCTCTTGTCTC 1029
QY 924 TTCTGCACTACCGTCTCAAAATCTGTGCGTAATGCAATACCACTGTGGCTTTTATCT 983
Db 1030 TTCTTGCAATACCGTATAACAAACCCAGTGCCTTTGCTCTATGGCTGGTCTCT 1089
QY 984 GTTATATGAGAGATTTGGTTGCTTTATCTGTGATCTGATCTGATCAGTTCCTCGAAGTGGTTT 1043
Db 1090 GTGATATGTGAGATCTGGTTTGGCTTATCTGGAATTTGGATCAGTTTCCCAAGTGGTTT 1149
QY 1044 CCAATCAACCGGAGACCTACCTTTGATAGACTGGCTTTTAAAGTATGACCCGAGAGGTGAA 1103
Db 1150 CTTGTGACCGGTGAACCTTACCTCGACAGCTTTCCTTTAAGATATGATCGTGAAGGTGAG 1209
QY 1104 CCGTCTCAGTTGGCTGTGTGTGACATATTTGTGATGACAGTCGACGCCCTTGAAGGAGCCA 1163
Db 1210 CCATCAGTTAGCTGTGTGTGACATTTTGTGATCTGCTGTGATTAATCCGTGGACAAG 1269
QY 1164 CCTATGCTACTGCCAACACTGTGCTATCCATTTCTGCTGTGATTAATCCGTGGACAAG 1223
Db 1270 CCCCCTGTGACGCAACACAGTGTCTCTATTTCTGGCTGTGATTAATCCAGTTGACAG 1329
QY 1224 GTCTTGTCTATGATCTGATGACGGAGCTTCAATGCTGACTTTTGAAGCAATGGCTGAG 1283
Db 1330 GTGCTCTGTATGTTCTGTGATGATGGTGTCTGTATGTTATCAATTTGAATCAGTGCAGAA 1389
QY 1284 ACTTCAGAGTTTGTAGAAATGGGTACCAATTTGTGGAAGAGTATGACATTTGAACCCAGA 1343
Db 1390 ACATCAGAGTTTGTCTGTAATGGGTACCAATTTTGAAGAAATATAGCAGAGCCCTCGT 1449
QY 1344 GCTCCCGAGTTTACTTTTCCAGAAATTTGATTTACCTGGAAGCAAGTCCAGCCTTCA 1403
Db 1450 GCACCAAGATTTGTTCTGTGCGAAATAGATTTACTTGAAGGATAAAGTTTCAGACATCA 1509
QY 1404 TTTGTTAAAGACCCCGGGCCATGAAGAGAAATATGAAGAAATTTAAATCAGGATAAAT 1463
Db 1510 TTTGTCAAGATCTGATAGCTATGAAGAGGAATATGAGGNAATTAATCCGAATCAAT 1569
QY 1464 GCCTAGTTTCTAAGGCAATGAAGTCCCGAGAGAGATGGATCATGCAAGATGGCACA 1523
Db 1570 GCATTTGTTCCAAAGCCCTAAAATGCTTGAAGAGGGTGGGTATGCAAGATGGCACA 1629
QY 1524 CCATGGCCAGGAACATACAGGATCATCTCGAATGATCAGGTTTCTCTGTCAC 1583
Db 1630 CCGTGGCTTGAATAATACAGGGGACCATCCAGGAATGATCCAGGTTCTTTAGGGCAA 1689
QY 1584 AGTGGTGGCTTGTATCTGAGGGTAATGAGCTCCCGGCTTTTGTATGTTCTCGTGAA 1643
Db 1690 AATGTTGACTTGTATGACAGAGGCAATGAGCTCCCGGCTTTTGTATATGTTCTCGAGAA 1749
QY 1644 AAGCGTCTGGGTTCCAGACCAACAGAGGCTGGTGCATGAATGCCCTTGTGCTGTC 1703
Db 1750 AAGCGACCGAGTTTCCAGCACCAACAAAGGCTGGTGTATGAATGCAATGCGTGAAGATT 1809
QY 1704 TCAGCTCTCTTACTAATGGAACATACATGTTGAATCTTGATTTGATGATCACTACATCAAC 1763
Db 1810 TCAGCAGTTCTTACCAATGGAACCTTTCACTTTGAATCTTGATTTGATGATCATTACATAAT 1869
QY 1764 AACAGCAAGGCTGTCCGAGAGAGCTATGTTGCTTCTTAATGGAATCCAAACCTAGGTCGCA 1823
Db 1870 AACAGCAAGGCTTAAAGAGAGCAATGCTTCTCTGTGAGACCCAAACCTCGGGAAGCAA 1929
QY 1824 GTCTGTTATGTCAGTGTCCCAAGAGTTTGAATGGATGATGAGGATGATGATGATGCA 1883
Db 1930 GTTTGTTATGTTGATGTTCCCAAGAGATTTGATGATGATGATGATGATGATGATGATGCT 1989
QY 1884 AACAGGAACACTGTCTTTTGTATTAATTAATTTGAGGGGCTTTGACGGCAATTAAGAGGACCA 1943
Db 1990 AATCGTAATACCGTGTCTTTGATATTAATTTGAGAGTTTAGATGGATTTCAAGGACCT 2049
QY 1944 GTTTATGTTGAACCTGGTTGTGTTTCAAGAGACAGCTATCTATGTTGTTATGAGCCCCCA 2003

2050 GTATATGTCGAACTGATGTGTTTTCACAGAAACAGCAATTAACGGTTATGAACCTCCA 2109
2004 ATTAAAGG-----CGAAGAGCCAGGTTTCTTGGCATCACTATGTCGGGCAAGAAAG 2057
2110 ATAAAGTAAACACAGAAAGCCAGTCTTTTATCTAAGCTCTGTGTGATCAAGAAAG 2169
2058 GCAAGCAAGTCAAGAAAGAGGAGCTCAGATAAGAAAGTCAAGCAAGCATGTGGACAGT 2117
2170 AAGAAATCCAAAGCTAAGAAAGAGTCGACAAAGAAATCAGGCAGGCATCTGACTCA 2229
2118 TCTGTTCCAGTATTTCAATCTCGAAGACATAGAGAGGGTGTGAAAGTCTGGTGTGAT 2177
2230 ACTGTTCTGTATTTCAACCTCGATGACATAGAAAGAGGGAGTTGAAGGTGCTGGTTGAT 2289
2178 GATGAGAAATCAGTCTCTCATGCTCAAAATGAGCTTAGAGAAAGAGATTTGGCCAGTCAGCA 2237
2290 GATGAAAGGCGCTCTTAATGTCGAAATGAGCTCGAAGAGCATTTGACAGTCTGCT 2349
2238 GCATTTGTTGCCCTCCACTCTGATGGAATATGGTGGTGTTCCTCAGTCTCTCCATCCAGAA 2297
2350 GTTTTGTGTTGCTTCTACCTTAATGGAATAATGGTGGTGTTCCTCCTTCAGCAACTCCAGAA 2409
2298 TCTCTTTGAAAGAGCTATCCATGTCATTAAGTTGTGGCTATGAGGACAACTCTGAATGG 2357
2410 AACTTTCTCAAAGAGGCTATCCATGTCAATTAGTTGTGGTTATGAGGATAAGTCAGATTGG 2469
2358 GGAACCTGAGATTTGTTGGATCTATGATCTGTCAAGAAATATTTCTTACTGATTTCAAG 2417
2470 GGAATGGAGATTTGATGATCTATGTTCTGTGACAGAAATATTTCTGACTGGTTCAAA 2529
2418 ATGACGCAAGAGGCTGGCGTTTCAGTCTAATTTGATGCCAAGCGCCAGCTTTTCAAGGGA 2477
2530 ATGCATGCCGCTGATGGCGATCCATTTACTGATGCTTAAGCTTTCAGCTTTCAGGGT 2589
2478 TCTGCCCCCATCAATCTTTGATGCTGTGAAACCAAGTCTGCGGTGGCTCTCGTCT 2537
2590 TCTGCTCTATCAATCTTTTCAGATCTGTGAACCAAGTCTGAGGTGGCTTTAGGTTCA 2649
2538 GTTGAATCTTTTCAGCGGCAATGCGCTTATGTTGATGCTACGAGAGGCGCTCAAG 2597
2650 GTTGAGATTTCTTTTCACTCGGCAATGCTCTATGTTGTTTACATGTTGATGAGGCTAATA 2709
2598 TTTCTGAGAGATTCGCTTACATCAACACCACTTTTACCACATTAACCTCTCTCCGCTT 2657
2710 TTTCTGAGAGGTTTGGTATGTAACACCACTATCTACCTATCACTCCATTCCTCTT 2769
2658 CTAGTCTATTTGATATTTGCTGCTATCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2717
2770 CTCATGTTATTTGATCTAGCGGTTGCTCTTCCACCAACCAAGTTTATTTATTCCTCAG 2829
2718 ATTAGCAACTTGGCCAGTATCTGTTCAATGCGCTCTTCTCTTCAATTTTTCGCCACTGGT 2777
2830 ATTATGAACATGCAAGTATATGTTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2889
2778 ATCTTTCAGATGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2837
2890 ATACTGAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2949
2838 TGGTCAATGAGGATCTCTGACATCTGTTTGGCGCTCTTTCAGGCTCTCTGAGGCTGCTGAGG 2897
2950 TGGTCAATGAGGATCTGCTCTATTTTGTCTGTTTTCAGGTTATCTCTCAAGATCTCTCAAG 3009
2898 CTTGCCGTTATCGACACCACTTCACTGTCACTCAAGGCTAATGACAGAGAGGCGAC 2957
3010 CTTGCCGTTATGACACCACTTCACTGTTTACCTCAAGGCTTCAAGATGAGACGGAGAC 3069
2958 TTTGCTGAGCTCTACATGTTCAAGTGGACGAGCTTCTCATCTCCCTCCGACGACATTTTG 3017
3070 TTTGCTGAGCTCTACTTGTTCAAATGGACCAACACTTCTGATTTCCGCAACGACCTGCTC 3129
3018 ATCAATTAACATGTTGTTGTTGCTGTCAGGCTCTCTACGCTCAACAGTGGTTTACCA 3077
3130 ATTGTAACTTAGTGGGAGTTGTTGACAGGAGTCTCTTATGCTATCAACAGTGGATACCA 3189

Db	1150	CCTGTGAACCGTGAACCTACCTCGACAGCGCTGCTTTAAGATATGATCGTGAAGGTGAG	1209	Qy	2178	GATGAGAAATCAGTTCTCATGTCTCAATCAGCTTAGAGAAGAGATTGGCCAGTCAGCA	2237
Qy	1104	CCGTCTCAGTTGCTGCTGTGACATATTTGTTCAGTACAGTCGACCCCTGAGGAGCCCA	1163	Db	2290	GATGAAAAAGCGCTCTTAATGTGCAAAATGAGCGCTGGAGAGAGGATTTGGACAGTCTGCT	2349
Db	1210	CCATCAGATTAGCTGCTGTGACATTTTCGTGAGTACTGTTCACCCCTTTGAAGGAGCCCA	1269	Qy	2238	GCATTTGTTGCCCTCCACTCTGATGGAAATATGTTGGTGTCTCCTCAGTCTCTCACTCCAGAA	2297
Qy	1164	CCTATCGTCTACTGCGCAACACTGTGCTATCCATTTCTTGCTGTGATTTATCCCGTGGACAAG	1223	Db	2350	GTTTTGTGTTGCTTCTACCTTAATGGAATAATGGTGGTGTCTCTCCTTTCAGCAACTCCAGAA	2409
Db	1270	CCCTTGTGACAGCCCAACAGAGTCTCTATTTCTGGCTGTTCACCTACCCAGTGTGACAAG	1329	Qy	2298	TCTCTTTTGAAGAAGCTATCCATGTCTATAAGTTGTGGCTATGAGACAGATCTGATGG	2357
Qy	1224	GTCTCTGTCTATGATCTATGATGACGGAGCTTCAATGCTGATCTTTTGACGCATTTGGCTGAG	1283	Db	2410	AACCTTTCTCAAGAGAGCTATCCATGTCTATAGTTGTGGTGTATGAGGATAAGTCAGATGG	2469
Db	1330	GTGCTCTGTATGTTTCTGATGATGCTGCTCTATTTCTGGCTGTTCACCTACCTTGAATCACTTGCAGAA	1389	Qy	2358	GGAACTGAGATTGGTTGGATCTATGGATCTGTCTACAGAAGATATTTCTTACTGATATCAAG	2417
Qy	1284	ACTTCAGAGTTTCTAGGAATGGGTACATTTTGTGAGAGAGTATGACATTTGAACCCAGA	1343	Db	2470	GGAAATGGAGATTGGATGGATCTATGGTCTGTGACAGAAGATATTTCTGACTGGTTCAAA	2529
Db	1390	ACATCAGAGTTTCTGCTGTAATGGGTACCAATTTTGCAGAGAAATATAGCATAGAGCCCTCGT	1449	Qy	2418	ATGACGCAAGAGCGCTGGCGTTTCAGTCTATTTGATGCCAAGCCGACGCTTTCAAGGGA	2477
Qy	1344	GCTCCCGAGTTTACTTTTCCAGAAATTTGATTACTCTGAAAGACAAAGTCCAGCCCTTCA	1403	Db	2530	ATGCATGCCCGTGGATGGGATCCATTTACTGATGCTTAAGCTTCAGCTTTCAAGGGT	2589
Db	1450	GCACCAAGATGTTACTTTGCTGCGAAATAGATTACTTGAAGATAAAGTTTCAGACATCA	1509	Qy	2478	TCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTCTCGGTGGGCTCTCGGTTCT	2537
Qy	1404	TTTGTAAAGACCGCCGGGCGCATGAAGAGAGATATGAAGAAATTTAAATTCAGATTAAT	1463	Db	2590	TCTGCTCTCTATCAATCTTTTCAGATCGTCTGAACCAAGTCTGAGGTGGGCTTTAGGTTCA	2649
Db	1510	TTTGTCAAAGATCCTAGAGCTATGAGAGGGAATATGAGGAATTTTAAATCCGAATCAAT	1569	Qy	2538	GTTGAAATCTTTTTCAGCGCGCATTCGCCCTTATGGTATGGCTACGAGAGGGCGCTCAAG	2597
Qy	1464	GCCTAGTTTCTAAGGCATTTGAAGTCCCGAGGAGGATGATCATGCAAGATGGACA	1523	Db	2650	GTTGAGATTCTCTTCAGTCGGCATTCCTATATGATGTTTACAATGGAGGCTAAAA	2709
Db	1570	GCATTTGTTTCCAAAGCCCTAAATGTCTTGAAGAGGGTGGTATGCAAGATGGACA	1629	Qy	2598	TTCCCTGGAGAGATTCGCTTACATCAACACCACTTTACCCATACTCTCTCCCGCTT	2657
Qy	1524	CCATGGCCAGGAACAATACAGGGATCATCTCGGAATGATTCAGGTTTTCTTGGTFCAC	1583	Db	2710	TTTCTTGGAGGTTTTCGTATGTGAACACCACTATCACTATCACTCCATTCCTCTT	2769
Db	1630	CCGTGGCCCTGAAAAATAACAGGGACCATCCAGGAATGATCCAGGTCTTCTTAGGGCAA	1689	Qy	2658	CTAGTCTATTGTATATTGCTGCTATCTGCTCTCACTGGAAGTTTCATCATGCCAGAG	2717
Qy	1584	AGTGGTGGCTTCTGATCTAGGGTATAGAGTCCCGCTTTAGTTATGTTGCTCTGTA	1643	Db	2770	CTCATGATTGTACATTTGCTAGCCGTTGTCTTTCACCAACAGTTTATTTCTCTCAG	2829
Db	1690	AATGGTGGACTTGTATGAGAGGGCAATGAGCTCCCGGTTTTGTATATGTTTCTCGAGAA	1749	Qy	2718	ATTAGCAACTTGGCCAGTATCTGGTTCAITGGCTCTTCTCTTCAATTTTCGCCACTGGT	2777
Qy	1644	AAGCTCTGGTTTCAGACCAACCAAGAGCTGGTCCCATGAATGCCCTTGTTCGTGTC	1703	Db	2830	ATTAGTAACATTGCAAGTATATGTTTCTGTCTCTCTTCTCTCTCTCTCTCTCTCTCT	2889
Db	1750	AAGGACCAAGATTCAGACCAACCAAGAGCTGGTGTCTATGATGACATCTGGTGAGATT	1809	Qy	2778	ATCCTTGGAGATGAGTGGAGTGGTGTGGCAATGACAGTGGTGGAGGATGAACAGTTC	2837
Qy	1704	TCAGCTGTCTTACTAATGGACAATACATGTTGAATCTTGAATGTGATCTACTACATCAAC	1763	Db	2890	ATATGAAATGAGTGGAGTGGCTGAGGATAGACGAATTTGGTGGAAACGAGCAGTTC	2949
Db	1810	TCAGCAGTTCTTACCAATGACCTTTTCATCTTGAATCTTGAATGTGATCATTAATAAT	1869	Qy	2838	TGGGTCACTGGAGGTATCTCTGCACATCTGTTTGGCTCTTTCAGAGGTCTCTGAAAGTG	2897
Qy	1764	AACAGCAAGCTGTCCGAGAGCTATGTTGCTTCTTAATGATCCAACTAGTTCGCCAA	1823	Db	2950	TGGGTCACTGGAGGTATCCGCTCATTTANTCGCTGTGTTTCAAGGTATCTCTCAAGTC	3009
Db	1870	AACAGCAAGCTTTAAGAGAGCAATGTGCTTCTGTATGGACCCAAACCTCGGGAAGCAA	1929	Qy	2898	CTTGGCGGTATCGACACCAACTTCCTCTCACCTCAAAGGTAAATGACGAAGAGCGCAC	2957
Qy	1824	GTCTGTTATGTCAGTTCCACAAAGTTTGTATGGGATTCATAGGAATGATCGATATGCA	1883	Db	3010	CTTGGCGGTATTCACACAACTTCACAGTTTACCTCAAAGCTTCAGATGAAGACGGAGAC	3069
Db	1930	GTTTGTATGTTCCAGTTCCCAAGAGATTGTATGTTGATTCGATGAAGACGATAGATGCT	1989	Qy	2958	TTTGTGAGCTCTACATGTTCAAGTGGAGCAGCGCTCTCATCTCTCGACGACCATTTTG	3017
Qy	1884	AACAGCAAGCTGTCTTTTGTATTAATTAACCTTGAAGGGCCCTTGACCGCATTCGAAGACCA	1943	Db	3070	TTTGTGAGCTCTACTTGTTCAAATGGACAACACTTCTGATTCGCGCAACGACGCTGCTC	3129
Db	1990	AATCGTAATACCGTGTCTTTGTATTAATTAACCTTGAAGGTTTGTAGATGGATTCGAAGACCT	2049	Qy	3018	ATCATTAACATGTTGGTGTGTTTGGCACCTCTTACGCCATCAAAGTGGTGTACAA	3077
Qy	1944	GTTTATGTGGAACTGTTGTGTTTCAACAGAACGCTATCTATGTTTATGAGTATGAGCCCCA	2003	Db	3130	ATTGTAACTTGTGGAGTGTGTGAGGAGTCTCTTATGCTATCAACAGTGAATACAA	3189
Db	2050	GTATATGTGGAACTGGATGTGTTTTCACAGAACAGCATTTATACGGTTTATGAACCTTCCA	2109	Qy	3078	TCATGGGGCGCTCTTTTGGGAAAGCTCTCTTTGCCCTTCTGGGTGATTTGTCTACTTATAC	3137
Qy	2004	ATTAAGG-----CGAAGAGACCAAGGTTTCTTGGCATCACTATGTGGGGGCAAGAAG	2057	Db	3190	TCATGGGGCACTCTTTTGGTAAAGTTGTTCTTTGCCCTTCTGGGTGATTTGTTCACTTGTAC	3249
Db	2110	ATAAAGTAAACACAGAGAGCCAAAGTCTTTTATCTAAGCTCTGTGGTGAATCAAGAAAG	2169	Qy	3138	CCATTCTCAAGGCTCTTATGGCAGGCAAAAACGACACAGGAGTGTGATCGTCTGG	3197
Qy	2058	GCAAGCAAGTCAAGAAAAAGAGCTCAGATAAGAAAAAGTCCGAACAGCATGTGGACAGT	2117	Db	3250	CCTTCTCAAGGTTTGTATGGTTCGACAGAACCGGACTCTCTACCATTTGTGTGGTCTGG	3309
Db	2170	AAGAAATCCAAAGCTAAAGAAAGAGTCCGACAAAGAAAGAAATCAGGCGAGCATACTGACTCA	2229	Qy	3198	GCTGTCTCTCGCTCTCATCT	3257
Qy	2118	TCTGTCTCAGTATTCATCTCGAGACATAGAGAGGTTTGAAGGTGCTGGGTTTGTAT	2177	Db	3310	TCGTGTTCTTGGCTTCTATCT	3369
Db	2230	ACTGTTCTCTGATTCAACCTCGATGACATAGAGAGGAGTTGAAGGTGCTGGTTTTGTAT	2289				

[illegible]

Search completed: August 23, 2004, 05:27:48
Job time : 13714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2004, 01:10:50 ; Search time 1303 Seconds

(without alignments)
11821.923 Million cell updates/sec

Title: US-09-900-237A-29

Perfect score: 3626

Sequence: 1 gcacgagaaacccgcctcca.....tcacatttggaggagtttt 3626

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N Geneseq 29Jan04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002s:*
- 7: geneseqn2003as:*
- 8: geneseqn2003bs:*
- 9: geneseqn2003cs:*
- 10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	2637.8	72.7	3264	7 ADA69779	Ada69779 Rice gene
2	2595	71.6	3799	6 AAS16458	Aas16458 Corn cDNA
3	2572.2	70.9	3746	3 AAZ299512	Aaz299512 DNA encod
4	2572.2	70.9	3746	3 AAZ299527	Aaz299527 DNA encod
5	2572.2	70.9	3773	3 AAZ299494	Aaz299494 DNA encod
6	2352.4	64.9	3222	7 ADA69499	Ada69499 Rice gene
7	2325.2	64.1	3704	3 AAZ299533	Aaz299533 DNA encod
8	1681.6	46.4	1734	3 AAZ58271	Aaz58271 Wheat cel
9	1677.6	46.3	3198	3 AAC49550	Aac49550 Arabidops
10	1673.6	46.2	3614	2 AAV06567	Aav06567 Arabidops
11	1408.4	38.8	3725	3 AAZ299500	Aaz299500 DNA encod
12	1408.4	38.8	3725	3 AAZ299521	Aaz299521 DNA encod
13	1408.4	38.8	3725	3 AAZ299506	Aaz299506 DNA encod
14	1382	38.1	3851	3 AAZ299506	Aaz299506 DNA encod
15	1373	37.9	3753	3 AAZ299530	Aaz299530 DNA encod
16	1373	37.9	3753	3 AAZ299515	Aaz299515 DNA encod
17	1373	37.9	3780	3 AAZ299497	Aaz299497 DNA encod
18	1343.4	37.0	3603	2 AAV06565	Aav06565 Arabidops
19	1340.2	37.0	3786	3 AAZ58265	Aaz58265 Corn cell
20	1340.2	37.0	3813	3 AAZ299509	Aaz299509 DNA encod
21	1340.2	37.0	3813	3 AAZ299524	Aaz299524 DNA encod
22	1338.6	36.9	3673	2 AAV06568	Aav06568 Arabidops
23	1331.4	36.7	2830	6 AAS16455	Aas16455 Corn cDNA

24	1320.4	36.4	3776	3 AAZ58263	Aaz58263 Corn cell
25	1317.2	36.3	3568	3 AAZ299491	Aaz299491 DNA encod
26	1271	35.1	3936	3 AAZ58266	Aaz58266 Corn cell
27	1271	35.1	3969	3 AAZ299518	Aaz299518 DNA encod
28	1271	35.1	3969	3 AAZ299503	Aaz299503 DNA encod
29	1257	34.7	3517	3 AAZ58268	Aaz58268 Soybean c
30	1217.2	33.6	3328	2 AAV34432	Aav34432 Cotton ce
31	1212.4	33.4	3207	2 AAV08372	AAV08372 Cellulose
32	1201.8	33.1	3747	3 AAZ671145	Aaz671145 Eucalyptu
33	1192.6	32.9	3255	6 ABZ12754	Abz12754 Arabidops
34	1191.4	32.9	3255	6 ABZ12754	Abz12754 Arabidops
35	1183.4	32.6	2890	4 AAZ58270	Aaz58270 Soybean c
36	1177	32.5	3444	4 AAC83798	Aac83798 Arabidops
37	1176.6	32.4	3828	2 AAV06566	Aav06566 Arabidops
38	1147.2	31.6	3311	2 AAV08373	AAV08373 Cellulose
39	1133.6	31.3	3232	4 AAC65448	Aac65448 Populus t
40	984.2	27.1	2248	2 AAV06562	Aav06562 Arabidops
41	966.8	26.7	2125	3 AAZ58269	Aaz58269 Soybean c
42	933.6	25.7	2033	2 AAV08381	AAV08381 Cellulose
43	921	25.4	2306	2 AAV34433	Aav34433 Cotton ce
44	863.4	23.8	1401	3 AAC44966	Aac44966 Arabidops
45	835.4	23.0	2055	3 AAA67111	Aaa67111 Eucalyptu

ALIGNMENTS

RESULT 1

ADA69779	ID	ADA69779	standard; DNA; 3264 BP.
XX	AC	ADA69779;	
XX	DT	20-NOV-2003 (first entry)	
XX	DE	Rice gene, SEQ ID 3102.	
XX	KW	Plant; bacterial infection; fungal infection; viral infection; rice;	
XX	OS	Oryza sativa.	
XX	PN	WO2003000898-A1.	
XX	PD	03-JAN-2003.	
XX	PF	22-JUN-2001; 2001WO-IB001105.	
XX	PR	22-JUN-2001; 2001WO-IB001105.	
XX	PA	(SYGN) SYNGENTA PARTICIPATIONS AG.	
XX	PI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y, Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G; WPI; 2003-175290/17.	
XX	PT	Identifying at least one gene involved in plant resistance or response to pathogenic infection for conferring resistance or tolerance to a plant to bacterial, fungal or viral infection by determining or detecting plant gene expression.	
XX	PS	Claim 6; SEQ ID NO 3102; 899pp; English.	
XX	CC	The present invention relates to a method (M1) for identifying genes involved in plant resistance or response to pathogenic infection. M1 comprises identifying a gene whose expression is significantly altered in the incompatible interaction of plant gene expression relative to expression of the gene in an uninfected plant, in a mutant plant that does not express a gene associated with response to pathogenic infection, or in a corresponding incompatible or compatible interaction. (M1) is useful for conferring resistance to resistance or tolerance to a plant to bacterial, fungal or viral infection. The present sequence was used to	

Db 2041 TCAACAAGCAGCGTGACAGTCTGTGCGACGTTTTCATCTTGAAGATATAGAGAGGCT 2100
QY 2157 GTTGAAGTGTGGTTCATGATGAGAAATCAGTTCTCATGCTCTCAATGAGCTTAGAG 2216
Db 2101 GTTGAAGTGTGGTTCATGATGAGAAATCAGTTCTCATGCTCTCAATGAGCTTAGAG 2160
QY 2217 AAGAGATTGGCCAGTCAGCAGCAATTTGTCCTCCACTCTGATGAAATATGTTGGTGT 2276
Db 2161 AAGAGATTGGCCAGTCAGCAGGTTTGTGCTCCACTCTCATGAAATATGTTGGTGT 2220
QY 2277 CTTAGTCTCCACTCCAGAACTCTTTTGAAGAAGCTATCATGCTATAAGTTGTGGC 2336
Db 2221 CTTCAATCTGCAACCCAGAACTCTTTTGAAGAAGCTATCATGCTATAAGTTGTGGC 2280
QY 2337 TATGAGGCAAGTCTGAATGGGGAAGTGA-----GATTGGTGGATC 2378
Db 2281 TATGAGGCAAGCCGAATGGGGAAGTGAAGTATGCAACTTGGCAATGTTGGTGGATC 2340
QY 2379 TATGATCTGTACAGAAGATATTTCTACTGGATTCAAGATGCAAGAGAGCTGGCGT 2438
Db 2341 TACGGTTCGTCAGAGAGATATTTCTACTGGATTCAAGATGCAAGAGAGCTGGAG 2400
QY 2439 TCAAGTCTATGATGCCAAGCCAGCTTTTCAAGGAGATCTGCCCCCAATCAATCTTCA 2498
Db 2401 TCAATCTACTGATGCCAAGCCAGCTTTTCAAGGAGATCTGCCCTCAATCAATCTTCA 2460
QY 2499 GATCTCTGAACCAAGTCTGGGTGGGCTCTCGTTCTGTTGAAATCTTTTCAGCCGG 2558
Db 2461 GATCGTCTTAACCAAGTCTTGGGTGGGCTTGGTCTGTTGAAATCTTTTCAGTGGC 2520
QY 2559 CATTCGCCCTTATGGTATGGCTACGAGGCGGCCCTCAAGTCTCTGGAGAGATTCGGTTAC 2618
Db 2521 CATTCGCCATATGATGAGGCTATGGAGGACGCTTAAGTTCTTGGAGAGATTCGCTAC 2580
QY 2619 ATCAACACCAATTTACCACTCTCTCCGCTTCTAGTCTATTGTTATATGGCT 2678
Db 2581 ATCAACACCAATTTATCCATGTGATCATGATCCGCTTCTCATATATCTGTTTGGCT 2640
QY 2679 GCTATCTGTCTGCTCACTGGAAGTTTCATGATCCAGAGATTAGCAACTTGGCCAGTATC 2738
Db 2641 GCTATCTGTGTTGCTCAGTGGGAAGTTCAATCCAGAGATTAGCAACTTGTCTAGTAT 2700
QY 2739 TGGTTCAATGCGCTCTCTCTTCAATTTTCGCACTGTTATCTTGAGATGAGTGGAGT 2798
Db 2701 TGGTTCAATCTCTCTCTTCAATTTTCGCACTGTTATCTTGAGATGAGTGGAGT 2760
QY 2799 GGTGTGGCATGACAGTGTGGAGGAATGAACAGTTCTGGGTCAATGAGATGATCT 2858
Db 2761 GGTGTGGCATGATGATGAGTGGAGGAATGAACAGTTCTGGGTCAATGAGATGATCT 2820
QY 2859 GCACATCTGTTTGGCGTCTTTCAGGCTCTTCTGAAGTGTCTGCGGTATGACACCCAC 2918
Db 2821 GGGCAATCTTTTGGCGTCTTTCAGGCTCTTCTGAAGTGTCTGCGGTATGACACCCAC 2880
QY 2919 TTCACTGTCACTCAAGCTTAATGCAAGAAGGCGATTTTCTGAGCTCTACATGTTTC 2978
Db 2881 TTCACTGTCACTCAAGCTTTCTGATGAAGATGCGACTTTGCTGAGCTCTACATGTTTC 2940
QY 2979 AAGTGGACAGCTTCTCATCCCTCCGACGACCATTTTGAATCAATTAACATGTTGGTGT 3038
Db 2941 AAGTGGACAGCTTCTCATCCACGACGACCATTTGATCAATTAACCTGGTGGTGT 3000
QY 3039 GTTGTGGCACCTCTAGCCATCAACAGTGTGTACCAATCATGGGCGCTCTTTGGG 3098
Db 3001 GTTGTGGTATCTCATACGATCAACAGCGCTTACCAATCAAGGAGCGCTCTTTGGC 3060
QY 3099 AAGCTCTCTTGTGCTTCTGGTGTGTTGTTCACTTATACCAATCTCAAGGCTCTTATG 3158
Db 3061 AAGCTCTCTTGTGCTTCTGGTGTGTTGTTCACTTATACCAATCTCAAGGCTCTTATG 3120
QY 3159 GGCAGGCAAAACCGCACCGACAGTGTGCTGCTGGGCTGTCTCTGCTGCTCTATC 3218
Db 3121 GGTGCGCAAAACCGCACTCCGACCATCGTTGTTGTTGTTGGGCAATCTCTCTGGCTTCGATC 3180

QY 3219 TTCTCTCTCTGCTGGTTCGTTGATCCATTCACTACCCGTTCTGCTGGCCCAATATC 3278
Db 3181 TTCTCATTTGCTGGTTCGATCGATCCATTCAACCCGTTCTACCCGCGCCAGATACC 3240
QY 3279 CAAACCTGTGGCATCAACTGTCTA 3301
Db 3241 CAAACATGTGGCATCAACTGTCTA 3263

RESULT 2
AAS16458
ID AAS16458 standard; cDNA; 3799 BP.
XX AAS16458;
AC AAS16458;
XX
DT 14-FEB-2002 (first entry)
XX
DE Corn cDNA encoding cellulose synthase Cqrae19/cesa-19.
XX
KW Corn; ss; cellulose synthase; Cdp45; cesa-3; Cqrae19; cesa-9;
KW stalk quality; improved stand; silage; pericarp; kernel hardening;
KW handling ability; transgenic plant.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 238..3799
FT /*tag= a
FT /product= "Cellulose synthase"

XX WO200179516-A2.
XX
XX
PD 25-OCT-2001.
XX
XX
PF 12-APR-2001; 2001WO-US011951.
XX
XX
PR 14-APR-2000; 2000US-00550483.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Dhugga KS, Helentjaris TG;
XX
XX WPI: 2002-041338/05.
XX P-PSDB; AAU10496.

New cellulose synthase polypeptides and polynucleotides, useful in improving stalk quality or silage, and in increasing concentration of cellulose in the pericarp, hardening the kernel for improved handling ability.

Claim 4; Page 80-85; 89pp; English.

The invention relates to isolated nucleic acids encoding two cellulose synthase proteins from corn, Cdp45 (cesa-3) and Cqrae19 (cesa-9). Also disclosed are a recombinant expression cassette comprising the polynucleotide (operably linked to a promoter) a host cell comprising the recombinant expression cassette and a transgenic plant comprising the recombinant expression cassette. The nucleic acid is useful in the improvement of stalk quality for improved stand or silage, and in the increased concentration of cellulose in the pericarp, hardening the kernel, and thus improving its handling ability. The nucleic acids may also be used as probes or amplification primers in the detection, quantification or isolation of gene transcripts, as probes in detecting deficiencies in the level of mRNA, for detecting gene mutations or allelic variants, for monitoring up regulation of expression or changes in enzyme activity in screening assays, for site directed mutagenesis, and in sense or antisense suppression of one or more genes in a host cell, tissue or plant. The polypeptides may be used in assays for enzyme agonists or antagonists, and as immunogen or antigen to obtain antibodies specifically immunoreactive with the protein. The present sequence encodes a corn cellulose synthase of the invention

36	CGGGTTGATCGCTCTGCGCGCGCATGACGGCGACGGGACGGCGCTGGAAGTCCGGAGG	95
214	CGGTTCTGGGGAGCTCGCCTGCCATGAGGGCGACGGCGCGCTGGAAGTCCGGAGG	273
96	CACGGGGCGGGGAGCTGTGCCAGATCTGCGCCGACGGCCCTGGGCACACAGCTTGGACGGC	155
274	COCGGGGAGGGCAGGTGTGCCAGATCTGCGCGGATGGCTGGGCGCATGACGGCGGAGGA	333
156	GACGCTTTCACGGCTCGCAGCTGTGCGCTTCCCGGTCTGCCGCGCCCTGCTACGAGAC	215
334	GACGCTTTCACGGCTCGCAGCTGTGCGGTTCCCGGTGTGCCGCGCCCTGCTACGAGTAC	393
216	GAGCGCAAGGAGGGACCCAGGCTGCTCCAGTCAAGACCAAGTACAAGCGCCACAGA	275
394	GAGCGCAAGGAGCGGCACCAAGCTGCCCGCAGTGCACAAACAGTACAAGCGCCACAG	453
276	GGGAGCCACGAGATCCGCGGGGAGGACGACACTGATGCCGATGATGGTAGTAC	335
454	GGGAGTCCAGCGATCCGAGGGGAGGACGACGATCTGATGCCGATGATGCTAGCGAC	513
336	TTCAACTACCTTGCATCTGGCACTGAGACACGAGCAGAGATTTGCTGACAGATCGGC	395
514	TTCAACTACCTTGCATCTGGCAATGACACACGAGCAGAGATTTGCTGACAGGATGCGC	573
396	AGCTGGCGCATGAACACACCGGGGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGC	455
574	AGCTGGCGCATGAATGCTGGGGGACGCGGGATGTTGGCCGCCCAAGTATGACAGTGGT	633
456	GAGATCGGCTCTCOAGTATGACAGTGGAGAGATCCCTAGGGGATAGCTCCCTTCAGTC	515
634	GAGATCGGCTCTTACCAAGTACGACAGTGGTGGATTCCTCGGGGATACATCCCGTCACTC	693
516	ACCAACAGCAGATGTCAAGAGAAATCCCTGGAGCTTCGCGTGCATCATCATGATGCC	575
694	ACTAACAGCAGATTTCCGGAGAAATCCCTGGTGTCTCCCTGACCATCATGATGCT	753
576	CCTACGGGAAACATCAGCAGACGTGCTCGGTTTCCCTATGTGAATCATCAAAATCCG	635
754	CCTACTGGGAAACATTTGGCAGGCGGCCCAATTCCTCTATATGAATCATTCATAATCCG	813
636	TCAAGGGAGTCTCCGGCAGTATTCGGAGTGTGCTCGAAGAGAGAGTGTGATGGCTGG	695
814	TCGAGGGAATCTCTGGTAGCGTTGGGAATGTTGCTCGAAGAGAGGGGTGATGGCTGG	873
696	AAAAATGAACGAGCAAAAGGTGCGATTCCTGACTAATGGGACAAAGCATTTGCTCCCTCT	755
874	AAAAATGAACGAGCAAAAGGAAACAAATCCCATGCAAGATGGCACAAGCATTTGCTCCCTCT	933
756	GAAAGTGGGAGCTACTGACATCATGATCTACTGATACAAACATGGAAGAGCGTTTA	815
934	GAGGCGCGGGGTGTTGGTGATATTGATGATCAACTGATTAACAACATGGAAGATGCCTTA	993
816	CTGAATGATGAACCTCGCCAGCTCTATCTAGAAAAGTCCCAATGCTTCTCCAAAATA	875
994	TTAAACGATGAACCTCGCCAGCTCTATCTAGAAAAGTTCACATTCCTTCTCCAGGATA	1053
876	AATCCCTCAGAAATGGTCAATGTTCTGGGTTGGTGTTCCTAAGCATCTTCTGACATAC	935
1054	AATCCCATACAGATGGTCAATGTTGTGCTAGATTTGTTTAAGCATCTTCTTGCACATAC	1113
936	CGTCTCACAATAATCTGTGGTAATGCATACCACTGTGGCTTTTATCTGTTATATGTGAG	995
1114	CGATCACAATAATCTGTGGTAATGCATACCACTGTGGCTTTTATCTGTTATATGTGAG	1173
996	ATTGGTTGCTTTATCTCGGTATACTGGATCAGTTCCTCGAAGTGTTCCTCAATCAACCGG	1055
1174	ATCTGTTTGTCTTTCTCGGTATTTGATCATCATGTTTCCAAAGTGTTCCTCAATCAACCGG	1233

XX	Sequence	3746 BP; 861 A; 976 C; 1006 G; 902 T; 0 U; 1 Other;	
SQ	Query Match	70.9%; Score 2572.2; DB 3; Length 3746;	
	Best Local Similarity	86.4%; Pred. No. 0;	
	Matches 2864; Conservative	1; Mismatches 439; Indels 9; Gaps 2;	
QY	4	CGAGGAACCCGCTCCAGCTCTCTCGTCCGTGCGGGTTGGATCGCTCGCGCCCATGG	63
DB	265	CCAGGCCCAAGGCTCCAGGCCAGCTCCCTCGACGTTTCTCGCGAGCTGCTTGCCTATGG	324
QY	64	ACGGCGACCGGACGGCCCTGAAGTCCGGGAGGCACCGGGCGGGGACGGTGTGCCAGATCT	123
DB	325	AGGGCGACCGGACGGCGTGAAGTCCGGGAGGCGCGTGGCGACAGGTTGTGCCAGATCT	384
QY	124	GGCCGACGGCTGGGACCAACGTTGGACGGCGACGTCCTTACCGCTCGGACGTCGTC	183
DB	385	GGGGCGACGGCGTGGGACCCACGGCGGAGGGGACGTCCTTCGCGGCTCGACGTCGTC	444
QY	184	GCTTCCGGTCTCCGCGCCCTGCTACGAGCACGAGCGCAAGGAGGCGACCCAGGCTGCC	243
DB	445	GGTTTCGGTGTGCGGCCCTGCTACGAGTACAGCGCAGGACGCGACGCGCGTGCC	504
QY	244	TCCAGTGCAGAACCAAGTACAAGCGCCACAGAGGAGCCGACGATCCGCGGGAGGAAG	303
DB	505	CCCAAGTGCAGAACCAAGTACAAGCGCCACAAAGGGAGCGCGCGATCCGTTGGGGAGGAAG	564
QY	304	GCGACACACTGATCGGATGATGGTAGTGACTTCAACTACCTGCTATCTGGCACTGAGG	363
DB	565	GAGACACACTGATGCCGA-----TAGCGACTTCAATTACCTTGCATCTGGCAATGAGG	618
QY	364	ACCAGAGCAGAAAGATTGCTGACAGGATGCGCAGCTGGCGCATGAACACCGGGGCGAGTG	423
DB	619	ACCAGAGCAGAAAGATTGCGCAGCAGATGGCGCAGCTGGCGCATGAACAGTTGGGGGAGCG	678
QY	424	GCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGSCCTCTCCAAGTATGACAGTG	483
DB	679	GGGATGTTGGTCCGCCCAAGTATGACAGTGGCGAGATCGSGCTTACCAAGTATGACAGTG	738
QY	484	GAGAGATCCCTAGGGATAGTCCCTTCACTACCAACAGCCAGATGTCAGGAGAAATCC	543
DB	739	GGAGATTCCTCGGGGATACATCCCATGCTACTAACAGCCAGATCTCAGGAGAAATCC	798
QY	544	CTGGAGCTTCGCTGATCATCATGATGTCCTTACGGGAAACATCAGCAGACGCTGCTC	603
DB	799	CTGGTCTTCCCTGACCATCATATGATGTCCCACTGGGAACATTGGCAAGCGTGTCTC	858
QY	604	CGTTTCCCTATGGAATCATTCACAAATCCGTCAGAGGAGTTCTCCGGCAGTATTGGGA	663
DB	859	CAPTTCCCTATGTGAACCAATTCCCAAAATCCGTCAGAGGAGTTCTCTGGTAGCATTTGGGA	918
QY	664	ATGTTCCCTGGAAGAGAGTTGATGGCTGGAAATGAGCAGGACAGGGTGGCATTC	723
DB	919	ATGTTCCCTGGAAGAGAGGTTGATGGCTGGAAATGAGCAGGACAGGGGACGATTC	978
QY	724	CCATGACTAATGGGACAAAGCATTCCTCCCTCTGAAGTCGGGACGCTACTGACATCGATG	783
DB	979	CCATGACGAATGGCACAGCATTTGCTCCCTCTGAGGGTCGGGGTGTGTTGATATTGATG	1038
QY	784	CATCTACTGAATACAACTGGAAGAGCGTTTACTGATGATGAATGAACCTCGCAGCCTCTAT	843
DB	1039	CATCAACTGATTACAACTGGAAGATGCTTATTGAACGACGAAATCTCGACAGCCTCTAT	1098
QY	844	CTAGAAAGTCCCATTTGCTTCTCCAAATAATCCCTACAGAATGGTCATTGTTCTGC	903
DB	1099	CTAGGAAGTTCACCTTCTTCTCCAGGAATAATCCATACAGATGGTCAITTTGCTGC	1158
QY	904	GSTTGTGTTCTTAAGCATCTTCTGCACTACCGTCTCAAAATCCTGTGCGTAATGCAT	963
DB	1159	GATTGATTGTTCTAAGCATCTTCTTGCATCTACCGTATCAAAATCCTGTGGCAATGCAT	1218
QY	964	ACCACCTGGCTTTTATCTGTTATATGTAGATTTGGTTTGTCTTTATCCTGGATCTGG	1023
DB			
1219	ACCAATTATGGCTTCTATCTGTTATATGTGAGATCTGGTTTGTCTTTCGTGATATTGG	1278	
QY	1024	ATCAGTTCCCGAAGTGGTTTCCAAATCAACCGGAGACCTACCTTGATAGACTCGCTTTAA	1083
DB	1279	ATCAGTTCCCTAAGTGGTTTCCAAATCAACCGGAGACCTACCTTGATAGACTCGCTTTAA	1338
QY	1084	GGTATCACCGAGAGGTGAACCGTCTCAGTTGGCTGCTGTTGACATATTTGTGAGTACAG	1143
DB	1339	GGTATGACCGGAGAGGTGAGCCATCTCAGTTGGCTGCTGTTGACATTTGTGAGTACAG	1398
QY	1144	TGACCCCTTGAAGGAGCCACTATCGTCACTGCCAACACTGTGCTATCAATTTCTGCTG	1203
DB	1399	TGACCCCAATGAAGGAGCCTCTCTTGTCACTGCCAATACCGTGTCTATCAATTTCTGCTG	1458
QY	1204	TTGATTTATCCGCTGGACAGGTTCTCTGCTATGATCTGATGACGAGCTTCAATGCTGA	1263
DB	1459	TGATTTATACCTGTGGATAAGGTTCTCTTGTATGATCTGATGATGAGCTGCGATCTGA	1518
QY	1264	CTTTTGACGCAATGGCTGAGACTTCAAGAGTTTCTAGGAAATGGGTACCATTTGTGAAGA	1323
DB	1519	CATTTGATGCACTAGCTGAGACTTCAAGAGTTTCTAGGAAATGGGTACCATTTGTGAAGA	1578
QY	1324	AGTATGACATTTGAAACCGAGACTCCGAGTTTCTTTTGGCCAGAAAATGATTAATCTGA	1383
DB	1579	AGTATGACATTTGAAACCGAGACTCCGAGTTTCTTTTGGCCAGAAAATGATTAATCTGA	1638
QY	1384	AGACAAAGTCCAGCCCTTCAATTTGTTAAAGACCGCGGCCCATGAAGAGATATGAAG	1443
DB	1639	AGACAAAGTCCAGCCCTTCAATTTGTTAAAGACCGCGGCCCATGAAGAGATATGAAG	1598
QY	1444	AATTTAAATTCAGGATAAATGCCCCTAGTTTCTTAAGGCAATGAAAGTCCCGGAGAAAGGAT	1503
DB	1699	AATTTAAATTCAGGATAAATGCCCCTAGTTTCTTAAGGCAATGAAAGTCCCGGAGAAAGGAT	1758
QY	1504	GGATATGCAAGATGGCAACCATGGCCAGGAAACAAATACCGAGGATTCATCTCGGAATGA	1563
DB	1759	GGATATGCAAGATGGCAACCATGGCCAGGAAACAAATACCGAGGATTCATCTCGGAATGA	1818
QY	1564	TTGAGTTTCTTCTGCTGACAGTGGTGGCTTGTATCTGAGGTAATGAGTCCCGCTT	1623
DB	1819	TTGAGTTTCTTCTGCTGACAGTGGTGGCTTGTATCTGAGGTAATGAGTCCCGCTT	1878
QY	1624	TAGTTTATGTGCTCGTGAAGCGCTCCTGGGTTCCAGCACCAAGAGGCTGGTGCCA	1683
DB	1879	TGCTCTATGTTTCTCGTGAAGCGTCCCTGGATTCAGCATCACAGAAAGCTGGTGCCA	1938
QY	1684	TGATGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1743
DB	1939	TGAATGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1998
QY	1744	ATTGTGATCACTACATCAACACAGCAAGGCTGTCGGAAGCTATGTGCTTCTTAATGG	1803
DB	1999	ATTGTGATCACTACATCAACACAGCAAGGCTGTCGGAAGCTATGTGCTTCTTAATGG	2058
QY	1804	ATCCAAACCTAGTCCGCAAGTCTGTTATGTGCGAGTTCACCAAGGTTTGTATGGGATTG	1863
DB	2059	ACCCTAACTAGGAGGAGTGTCTGCTACGTCAGTTTCCCGACAGATTCGATGGCATTTG	2118
QY	1864	ATAGGAATGATCGATGCAAAACAGGAACTGCTCTTTTGTGATATTAATCTGAGGGGCC	1923
DB	2119	ACAGGAATGATCGATGCAAAACAGGAACTGCTCTTTTGTGATATTAATCTGAGGGGCC	2178
QY	1924	TTGACGGCAATCAAGGACCAAGTTTATGTGGGAACCTGGTTGTGTTTCAACAGAACAGCTA	1983
DB	2179	TTGATGGCAATCAAGGACCAAGTTTATGTGGGAACCTGGTTGTGTTTCAACAGAACAGCTC	2238
QY	1984	TCTATGTTTATGAGCCCCCAATTAAGCGGAAGCAAGCGTGTCTTGTGGATCACTATGTG	2043
DB	2239	TATATGTTTATGAGCCCCCAATTAAGCGGAAGCAAGCGTGTCTTGTGTCACTACTATGTG	2298
QY	2044	GGGGCAGAGAGGCAAGCAAGCTCAAGAAAGGAGCTCAGATAGAAAGAAAGCTCGAACA	2103
DB	2299	GGGGTAGGAGAGGCAAGCAAGCTCAAGAAAGGAGCTCAGATAGAAAGAAAGCTCGAACA	2355

QY 2104 AGCATGTGACAGTCTCTGTTCCAGTATTCAATCTCGAAGACATAGAGGAGGTGTTGAG 2163
Db 2356 AGCATGTGACAGTCTCTGTTCCAGTATTCAACCTTGAAGATATAGAGGAGGTGTTGAG 2415
QY 2164 GTGCTGGGTTTATGATGAGAAATCAGTTCTCTCATGTCTCAATGAGCTTAGAGAGAGAT 2223
Db 2416 GGCTGGATTGACGACGAGAAATCATTCTTATGTCATAATGAGCCTGGAGAGAGAT 2475
QY 2224 TTGGCCAGTCAGCAGATTTGTTGGCTCCACTCTCATGATGGAATATGTTGTTCTCAGT 2283
Db 2476 TTGGCCAGTCAGCAGATTTGTTGGCTCCACTCTCATGATGGAATATGTTGTTCTCAGT 2535
QY 2284 CTTCCACTCCAGAACTCTTTTGAAGAGCTATCCATCTCATAGTTTGTGGCTATGAGG 2343
Db 2536 CCGCACTCCGAGTCTCTTTGAAGAGCTATCCATCTCATAGTTTGTGGCTATGAGG 2595
QY 2344 ACAAGTCTGAATGGGAACTGAGATTTGTTGATCTATGGAATCTGTCAAGAGATATTC 2403
Db 2596 ACAAGTCTGAATGGGAACTGAGATCGGTGATCTACGGTTCTGTGACAGAGACATTC 2655
QY 2404 TTACTGGAATCAAGATGACGAGAGCTGCGGTTTCACTTATGCAATGCCAAGCGCC 2463
Db 2656 TCACCGGATTCAGATGACGCGGAGGCTGGCGGTGATCTACTGATGCCCAAGCGGC 2715
QY 2464 CAGCTTCAAGGGATCTGCCCCCATCAATCTTTCAGATCGTCTGAACCAAGTCTCGGT 2523
Db 2716 CAGCTTCAAGGGATCTGCCCCCATCAATCTTTCAGACCGTCTGAACCAAGTCTCGGT 2775
QY 2524 GGGCTCTCGGTTCTGTGAAATCTTTTTCAGCGGCAATGCCCCCTTATGTTATGCTACG 2583
Db 2776 GGGCTCTGGTTCGGTGGAGATCTTTCAGCGGCACTGCCCTCTGTGTAGGCTACG 2835
QY 2584 GAGGCGGCTCAAGTTCCTGGAGATTCGCTTATCAACCAACCACTTATACCACTAA 2643
Db 2836 GAGGCGGCTCAAGTTCCTGGAGATTCGCTTATCAACCAACCACTTATACCACTAA 2895
QY 2644 CTTCTCTCCGCTCTAGTCTATTATGTTATGTTATGCTGCTATGCTGCTCACTGAAAGT 2703
Db 2896 CTTCTCTCCGCTCTAGTCTATTATGTTATGTTATGCTGCTATGCTGCTCACTGAAAGT 2955
QY 2704 TCATCATGCGAGATTAAGCAATTTGGGCAAGTATCTGTTTCAATGGGCTCTTCTTTCAA 2763
Db 2956 TCATCATGCGAGATTAAGCAATTTGGGCAAGTATCTGTTTCAATGGGCTCTTCTTTCAA 3015
QY 2764 TTTTCCGACATCGTATCTTGAATGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAG 2823
Db 3016 TTTTCCGACATCGTATCTTGAATGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAG 3075
QY 2824 GGAATGACAGTCTGGGTCATTGGAGGTATCTCTGCACATCTGTTTGGCTTCTTACG 2883
Db 3076 GGAATGACAGTCTGGGTCATTGGAGGTATCTCTGCACATCTGTTTGGCTTCTTACG 3135
QY 2884 GTCTTCTGAAGTGCTTCCCGTATCGACCAACTTCACTGTCACTCAAAAGGCTAATG 2943
Db 3136 GCCTGCTCAAGTGTGGCGGCACTCGACCAACTTCACTGTCACTCAAAAGGCTAATG 3195
QY 2944 AGGAGAGGCACTTGTGAGTCTACATGTTCAAGTGAGAGAGGCTTCTATCTTCTCCTC 3003
Db 3196 AGGAGAGGCACTTGTGAGTCTACATGTTCAAGTGAGAGAGGCTTCTATCTTCTCCTC 3255
QY 3004 CGACGACCATTTGATCATTAACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3063
Db 3256 CGACGACCATTTGATCATTAACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3315
QY 3064 ACAGTGGTTTACCAATCATGGGGGCGCTCTTTTGGGAAGCTCTTCTTGGCTTCTGGTGA 3123
Db 3316 ACAGTGGTTTACCAATCATGGGGGCGCTCTTTTGGGAAGCTCTTCTTGGCTTCTGGTGA 3375
QY 3124 TTGTTCACTTATACCATTTCTCAAGGGTCTTATGGGAGGCAAAACCGCACCGACGA 3183
Db 3376 TCGTCCAGTCTACCGGTTCTCAAGGGCTCATGGGAGGAGGAGGAGGAGGAGGAGGAGG 3435

QY 3184 TTGTCATCGTCTGGGCTGTCTCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3243
Db 3436 TGTGCTGTCTGGGCTGTCTCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCG 3495
QY 3244 ATCCATTCACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTAC 3303
Db 3496 ACCCTTCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTAC 3555
QY 3304 AAAGTGGAGTTT 3316
Db 3556 GAAAGTGAAGTT 3568

RESULT 4
AAZ99527
ID AAZ99527 standard; DNA; 3746 BP.
XX AAZ99527;
XX AC AAZ99527;
XX XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 321..3449
FT /*tag= a
FT /product= "cellulose synthase"
FT /transl_except= (pos: 1800..1802, aa: Xaa)
FT /note= "no termination codon given; Xaa is an unspecified
amino acid"
XX
PN WO200009706-A2.
XX
XX 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
PR 17-AUG-1998; 98US-0096822P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI: 2000-224343/19.
DR P-PSDB; AAY84119.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
XX
PS Claim 1; Page 176-181; 119pp; English.
XX
CC The present sequence encodes a maize cellulose synthase polypeptide. The
CC cellulose synthase can be used for the improvement of stalk quality for
CC improved stand or silage. It also provides an increased concentration of
CC cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants and
CC seeds expressing the cellulose synthase. The polynucleotide is used for
CC modulating, preferably increasing, the level of the synthase in a plant
CC cell. The plants are preferably monocots. The polynucleotide is also used
CC as a probe or primer in the detection quantitation or isolation of gene
CC transcripts. The probes are useful in detecting deficiencies in the level
CC of mRNA in screenings for desired transgenic plant, for detecting or
CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of

CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein

XX SQ Sequence 3746 BP; 861 A; 976 C; 1006 G; 902 T; 0 U; 1 Other;

Query Match 70.9%; Score 2572.2; DB 3; Length 3746;

Best Local Similarity 86.4%; Pred. No. 0;

Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2;

QY	4	CGAGGAACCCGCTCCAGCTCTCTCGTTCGGTGGGTGGATCGCTCTGCGCGCCATGG	63
DB	265	CGAGGCCCAAGGCTCCAGGCCAGCTCCTCGAGCTTCTCGGCGAGCTCGCTTGGCCATGG	324
QY	64	ACGGCGACCGGACGGCCCTGAAGTCCGGGAGGACCGGGCGGGGACGTTGTGCCAGATCT	123
DB	325	AGGGCGACCGGACGGCGTGAAGTCCGGGAGGCGCGTGGCGGACAGGTTGCGCAGATCT	384
QY	124	GGCGCGACCGCTGGGACCGACCTTGGACGGCGACGCTTTCACGGCTCGGAGCTGCC	193
DB	385	GGCGCGACCGCTGGGACCGACCGGCGGAGGCGGACGCTTTCGCGCGCTCGGAGCTGCC	444
QY	184	GCTTCCCGGTCTCGCGCCCTGCTACGACGACGAGCGCAAGGAGGCGACCGGCTGCC	243
DB	445	GGTTCCGGTGTGCGCGCCCTGCTACGAGTACGAGTACGAGTACGAGCGGCGGAGCGG	504
QY	244	TCAGTGCAGAACCAAGTACAAGCGCCACAGAGGAGGCCAGCGATCCGCGGGGAGGAAG	303
DB	505	CCGAGTGCAGAACCAAGTACAAGCGCCACAGGAGGAGCGCGGCGATCCGTTGGGAGGAAG	564
QY	304	GGACGACACTGATCGGATGATGCTAGTACCTTCACTTCACTTCACTTCACTTCACTTCACT	363
DB	565	GAGACGACACTGATGCGCA-----TAGCGACTTCAATTTACCTTGTCACTTGGCAATGAGG	618
QY	364	ACCAGAGCAGAGATGCTGACAGGATGCGAGCTGGCGCATGAACACCGGGGCGAGTG	423
DB	619	ACCAGAGCAGAGATGCTGCGACAGATGCGAGCTGGCGCATGAACAGTGGGGGCGAGTG	678
QY	424	GCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCGCTCTCCAAATATGACAGTG	483
DB	679	GGGATGTTGGTCCCGCCCAAGTATGACAGTGGCGAGATCGGCGCTTACCAAGTATGACAGTG	738
QY	484	GAGAGATCCCTAGGGGATAGTCCCTTCACTACCAACCGCAGATGTCAGGAGAAATCC	543
DB	739	GCAGATTCCTCGGGGATATATCCCATCACTCACTCACTCACTCACTCACTCACTCACTCACT	798
QY	544	CTGGAGCTTCGCTGATCATCATGATGTCCTTACCGGGAACATCAGCAGACGTCGCTC	603
DB	799	CTGGTCTTCCTCGACCATATATGATGTCCTTACCGGGAACATTGGCAAGCGTGTCTC	858
QY	604	CGTTCCCTATGTAATCATCAACCAATCCGTCAGGGAGTTCCTCGGCGAGTATGGGA	663
DB	859	CAITTCCTATGTAACCAATTCGCAAAATCCGTCAAGGGAGTTCCTCGTAGCATTTGGGA	918
QY	664	ATCTTCCTCGGAAGAGAGTTCGCTGGAATGAGAGCAGGACAGGGTGGCATTC	723
DB	919	ATGTTCCTCGGAAGAGAGTTCGCTGGAATGAGAGCAGGACAGGGGACGATTC	978
QY	724	CCATGACTAATGGGACAGCAATGCTCCCTCTCAAGGTTCGGGAGCTTACTGACATCGATG	783
DB	979	CCATGAGATGGCAGCAAGATGCTCCCTCTGAGGTTCGGGTGTTGGTATATGATG	1038
QY	784	CATCTACTGAATCAACATCGAAGACGCTTTACTGAATGATGAACCTCGCCAGCTCTAT	843
DB	1039	CATCAACTGATTAACAATCGAAGATGCTTTATTTGAACGAGAACTTCGACAGCTCTAT	1098
QY	844	CTAGAAAGTCCCATTTCTTCCCAAAATTAATCCCTACAGAACTGGTTCATTTGTTCTGC	903
DB	1099	CTAGAAAGTCCCATTTCTTCTCCAGGATTAATTCATACAGAGTGGTTCATTTGTTCTGC	1158
QY	904	GGTTGGTGTCTAAGCATCTTCTGCACTACCGTCTCAAAATCTCTGTCGTAATGCAT	963

DB	1159	GATTGATGTTCTAAGCACTCTTCTGCACTACCGTATCAAAATCTCTGTGCGCAATGCAT	1218
QY	964	ACCACTGTGGCTTTTATCTGTATATGTAGATTTGGTTTGTCTTATCTCTGATGACTGG	1023
DB	1219	ACCAATATGGCTTCTATCTGTATATGTAGATTTGGTTTGTCTTCTCTGTGATATGG	1278
QY	1024	ATCAGTTCCCGAAGTGGTTTCCAAATCAACCGGAGACCTTACCTTGTATAGACTCGCTTTAA	1083
DB	1279	ATCAGTTCCCTAAGTGGTTTCCAAATCAACCGGAGACCTTACCTTGTATAGCTGGCATTA	1338
QY	1084	GGTATCACCGAGAGGTGAACCGTCTCAGTTGGCTGCTGTGTGACATATTTGTGACAG	1143
DB	1339	GGTATCACCGGAGAGGTGAGCCATCTCAGTTGGCTGCTGTGTGACATATTTGTGACAG	1398
QY	1144	TCGACCCCTTGAAGGAGCCACCTATCGTCACTGCCAACAATGTGCTTCACTTCTTGGCTG	1203
DB	1399	TCGACCAATGAAGGAGCCCTCTCTTGTCACTGCCAATACCGTGTCTATCACTTCTTGGCTG	1458
QY	1204	TTGATTATCCCGTGGGACAAAGTCTCTTGTCTATGTATCTGATGACGAGCTTCAATGCTGA	1263
DB	1459	TGGATTTACCTGTGGATAGGTTCTTGTCTATGTATCTGATGATGAGCTGCGATGCTGA	1518
QY	1264	CTTTTACCGCATTTGGCTGAGACTTTCAGAGTTTGTCTAGGAAATCGGTAACCAATTTGTGAAGA	1323
DB	1519	CATTTGATGACCTAGCTGAGACTTTCAGAGTTTGTCTAGGAAATCGGTAACCAATTTGTGAAGA	1578
QY	1324	AGTATGACATTTGAACCCAGACTCCCGAGTTTACTTTTCCAGAAATTTGATTAACCTGA	1383
DB	1579	AGTATGACATTTGAACCCAGACTCCCGAGTTTACTTTTCCAGAAATTTGATTAACCTGA	1638
QY	1384	AAGACAAAGTCCAGCTTCTTCTTAAAGACCGCGGGCCATGAAGAGAGAAATATGAAG	1443
DB	1639	AGGACAAAGTGCACCTTCTTCTTAAAGACCGCGGGCCATGAAGAGAGAAATATGAAG	1698
QY	1444	AATTTAAATTCAGGATAAATGCCCCTAGTTTCTTAAGCATTTGAAAGTCCCGAGGAAGGAT	1503
DB	1699	AATTTAAATTCAGGATAAATGCCCCTAGTTTCTTAAGCATTTGAAAGTCCCGAGGAAGGAT	1758
QY	1504	GGATCATGACAGATGACACCATGCGCCAGGAACAATACCAAGGATTCATCTCTGGGAATGA	1563
DB	1759	GGATCATGACAGATGACACCATGCGCCAGGAACAATACCAAGGATTCATCTCTGGGAATGA	1818
QY	1564	TTGAGTGTTCCTTGTGTCACAGTGGTGGCTTCTGATCTGAGGGTAAATGAGTCCCGCGTT	1623
DB	1819	TTGAGTGTTCCTTGTGTCACAGTGGTGGCTTCTGATCTGAGGGCAATGAGTACCCCGTT	1878
QY	1624	TAGTTTATGTTCTCGTGAAGAGCGTCTGCGGTTCCAGCACCAAGAGGCTGGTGCCA	1683
DB	1879	TGCTCTATGTTTCTCGTGAAGAGCGTCTGCGGTTCCAGCACCAAGAGGCTGGTGCCA	1938
QY	1684	TGAATGCCCTTGTTCGTGTCAGCTGCTCTTACTTAATGAGCAATACATGTTGAATCTTG	1743
DB	1939	TGAATGCCCTTGTTCGTGTCAGCTGCTCTTACTTAATGAGCAATACATGTTGAATCTTG	1998
QY	1744	ATTGTGATCACTACATCAACACAGCAAGGCTCTCGGAGAGCTATGTGCTTCTCTATATGG	1803
DB	1999	ATTGTGATCACTACATCAACACAGTAAGGCTCTCAGGAGAGCTATGTGCTTCTCTATATGG	2058
QY	1804	ATCCAAACCTTAGTCCGCAAGTCTGTTATGTGAGTTTCCCAAAAGGTTTGTGAGGATGG	1863
DB	2059	ACCTTAACCTTAGGAGAGGAGTCTGCTACGTCGATTTCCCGAGAGATTCGATGGCATGG	2118
QY	1864	ATAGGATGATCATATGCAACAGGACACTCTCTTTTGTATATTAATCTCAGGGGCC	1923
DB	2119	ACAGGAATGATCATATGCAACAGGACACTCTCTTTTGTATATTAATCTCAGAGGTC	2178
QY	1924	TTGACGGCATTCAGGACCAAGTTTATGTGGAACTGTTGTGTTTCAACAGAACAGCTA	1983
DB	2179	TTGATGGCATTCAGGACCAAGTTTATGTGGAACTGTTGTGTTTCAACCGAACAGCTC	2238
QY	1984	TCATGTTTATGAGCCCCCAATTAAGCGAGAGGAGCCAGGTTTCTTGGCATCACTATGTC	2043

2239	TATATGGTTATGAGCCCCCAATTAAGCAGAAGAGGGTGGTTTCTTGTCACTACTATGGT	2299
2044	GGGCAAGAAGCAAGCAAGTCAAGAAAGAAAGGAGCTCAGATAAGAAAAGTGCAGACA	2103
2299	GCGETAGGAAGAGCAAGCAAAATCAAGAA--GGGCTCGGACAAGAGTTCGCAGA	2355
2104	AGCATGTGGACAGTTCTGTTCACGTATTCAATCTCGAAGACATAGAGAGGGGTGTTGAAG	2163
2356	AGCATGTGGACAGTTCTGTGCAGATTCAACCTTGAAGATATAGAGAGGGAGTTGAAG	2415
2164	GTGCTGGTTTGATGATGAGAAATCAGTCTCATGTCTCAATGAGCTTATAGAGAGAGAT	2223
2416	CGCTTGGATTGACGACGAGAAATCACTCTTATGTCTCAATGAGCCTCGAGAAAGAGAT	2475
2224	TTGGCCAGTCAGCAGCAATTTGTTGCCCTCCACTCTGATGGAATATGGTGGTCTCCTCAGT	2283
2476	TTGSCCAGTCCGACAGCGTTGTTGSCCTCCACTCTGATGGAGTATGGTGGTCTCCTCAGT	2535
2284	CCTCCACTCGAAMATCTCTTTTGAAGAAGATATCCATGTCTATAAGTTGTGGCTATGAGG	2343
2536	CCGCAATCTCGGAGTCTCTTCTGAAGAAGATATCCATGTTATAAGCTGTGGCTATGAGG	2595
2344	ACAACTCTGAATGGGAACTCAGATTCGTTGGATCTATGATCTGTCTCACAGAGATATTC	2403
2596	ACAAGACTGAATGGGAACTCAGATCGGGTGGATCTACGGTTCGTGACAGAGACATTC	2655
2404	TTACTGATTCAAGATCAGCAGAGAGGCTGGCGTTCAGTCTATTGCAATGCCAACAGCGCC	2463
2656	TCACCGGATTCAAGATCAGCGGAGGCTGGCGTTCGATCTACTGATGCCCAAGCGGC	2715
2464	CAGCTTTCAAGGGATCTGCCCCCAATCAATCTTTCAGATTCGTCTGAAACAAGTGTGCGGT	2523
2716	CAGCTTTCAAGGGTCTGCCCCCAATCAATCTTTGCGACCGTCTGAAACAAGTGTGCGGT	2775
2524	GGCTCTCGGTTCTGTTGMAATCTTTTCAGCGCGAATGCCCTTATGTTATGGCTACG	2583
2776	GGGCTCTTGGGTCCGTGGAGATCCCTTCACGCGGCACTGCCCGCTGTGTACGGCTACG	2835
2584	GAGGGCGCCTCAAGTTCCTGGAGAGATTCGTTACATCAACACCAACCATTTACCCACTAA	2643
2836	GAGGGCGGCTCAAGTTCCTGGAGAGATTCGGGTACATCAACACCAACCATTTACCCGTC	2895
2644	CCTCTCTCCGCTTCTAGTCTATTGTATATTGCCCTGCTATCTGTCTCTCACTGGAAAGT	2703
2896	CGTCCATCCCGCTTCTCATCTACTGCACTCCTGCCCGCACTGTGTCTCTCACCGGAAAGT	2955
2704	TCATCATGCAGAGATTAGCAACTTTGCGCAGATCTCGTTTCATTGCGCTCTTCTTTCAA	2763
2956	TCATCATTCAGAGATCAGCAACTTCGCCAGCATCTGGTTTCATCTCCTCTTTCATCTCGA	3015
2764	TTTTTCGCCACTGGTATCCTTTGAGATGAGTGGAGTGGTGTGGCATTTGAGAGTGGTGA	2823
3016	TCITTGCCACGGGCATCCTGGAGATGAGTGGAGCGGGTGGGCATCGACGATGGTGGGA	3075
2824	GGAAATGAACAGTTCTGGGTCAATTGAGAGATCTCTGCAACATCTGTTGGCGTCTTTCCAG	2883
3076	GGAAACGAGCAGTTCTGGGTGATCGGGGCATCTCCGGGCACCTCTTCGCGGTGTTCCAGG	3135
2884	GTCTTCTGAGGTGCTTGCCGATTCACACCAACTTCATCTGTACACTCAAGGCTAATG	2943
3136	GCCTGTCTCAAGGTGCTGGCGGCATCGACACCAACTTCACCGTCACCTCCCAAGGCTCTCG	3195
2944	ACGAAGAAGGCGACTTTGCTGAGCTCTACATGTTCAAGTGCAGACGCTTCTCATCCCTC	3003
3196	ACGAGGACGGGCATCTCGGAGGCTGTACATGTTCAAGTGGACACGCTCTGTATCCCGC	3255
3004	CGACGACCATTTTGATCAATTAACATGGTGTGGTGTGTTGCTGGCACCTCTCTACGCCATCA	3063
3256	CCACCAACATCCTGATCATCAACTGGTGTGGCGTTCGTCCGCGCATCTCCTACGCCATCA	3315
3064	ACAGTGGTTACCAATCATTGGGGCGGCTCTTTGGGAAGCTCTTCTTTGCTCTTGGGTGA	3123
3316	ACAGCGGATACAGTCTGTGGGGCGGCTCTTTCGGAAGCTCTTCTTCGCTCTTGGGTGA	3375

[illegible]

RESULT 5

AAZ99494
ID AAZ99494 standard; DNA; 3773 BP.

AAZ99494:

DT 03-JUL-2000 (first entry)

DE DNA encoding a maize cellulose synthase.

Maize; cellulose synthase; stalk quality; stand; silage; cellulose
transgenic plant; plant breeding marker; ss.
KW

OS Zea mays.

AA	Key	Location/Qualifiers
FH		
FT	CDS	337 3565

```

E1:  *tag=
E2:  /product=
E3:  "cellulose synthase"

```

```
/transl_except= (pos: 1817. .1819, aa: Xaa)
```

FT amino acid"

PN WO200009706-A2.

PD 24-FEB-2000.

16-AUG-1999: 99WO-US018760.

PR 17-AUG-1998: 98US-00968222P.

PA (PION-) PIONEER HI-BRED INT INC.

PI Dhugqa KS, Helentjaris TG, Bowen BA, Wang X; AA

WPI: 2000-224343/19.

UR
XX
F-PSDB; HAI84108.

PT new genes which encode maize cellulose synthase polyketides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.

PS Claim 1; Page 90-94; 119pp: English.

CC The present sequence encodes a maize cellulose synthase polypeptide. The
CC cellulose synthase can be used for the improvement of stalk quality for
CC improved stand or silage. It also provides an increased concentration of
CC cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants and
CC seeds expressing the cellulose synthase. The polynucleotide is used for
CC modulating, preferably increasing, the level of the synthase in a plant
CC cell. The plants are preferably monocots. The polynucleotide is also used
CC as a probe or primer in the detection quantitation or isolation of gene
CC transcripts. The probes are useful in detecting deficiencies in the level
CC of mRNA in screenings for desired transgenic plant, for detecting

CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein

XX Sequence 3773 BP; 863 A; 988 C; 1017 G; 904 T; 0 U; 1 Other;

Query Match 70.9%; Score 2572.2; DB 3; Length 3773;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2;

QY 4 CGAGGAACCCGCTCCAGCTCTGTGCTGCTGCGGTGGATGCTCTCGCGCCCATGG 63
DB 282 CCAGGCCAGGCTCCAGGCCAGCTCCCTCGACGTTTCTCGGCGAGCTGCTTGCCATGG 341

QY 64 ACGGCGACCGGACGCCCTGAAGTCGGGAGGACCGGGCCGGGAGCTGTGCCATCT 123
DB 342 AGGGCGACCGGACGCCGCTGAAGTCGGGAGGCGCGGTGGCGACAGGTGTGCCATCT 401

QY 124 GGGCCGACGCGCTGGGACACACGTTGCGACGGGAGCTTTCACCGCTCGGAGCTGCG 183
DB 402 GGGCGACGCGCTGGGACACACGCGGAGGGGAGCTTTCGCGCGCTCGGAGCTGCG 461

QY 184 GCTTCCCGTCTGCGCCCTCTGTACGAGCAGAGCGCAAGGAGGCGCCAGGCTGCG 243
DB 462 GGTTCGCGTGTGCGCCCTCTGTACGAGTACGAGCGAAGGACGCGACGAGCGTGGC 521

QY 244 TCCAGTGCAGACCAAGTACAGCGCCACAGGGGAGCCCGAGCTCCGGGGGAGGAG 303
DB 522 CCCAGTGCAGACCAAGTACAGCGCCACAGGGGAGCCCGGAGCTCCGTTGGGGAGGAG 581

QY 304 GCGACGACACTGATGCGCGATGATGTAGTACTTCAACTACCTGCTGCGACCTGAGG 363
DB 582 GAGCGACACTGATGCCG-----TAGCGACTTCAATTACCTTGGATCTGGCAATGAGG 635

QY 364 ACCAGAAGCAGAAGATTGCTGACAGGATGCGAGCTGGCGCATGAACACCGGGGCGAGT 423
DB 636 ACCAGAAGCAGAAGATTGCGGACAGAAATGCGAGCTGGCGCATGAACCTTGGGGGCGAG 695

QY 424 GCATTTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTCTCCAAGTATGACAGT 483
DB 696 GGGATGTTGGTCCGCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAAGTATGACAGT 755

QY 484 GAGAGATCCCTAGGGGATAGCTCCCTTCAGTCAACCAAGCCAGCATGTGAGGAGAAATCC 543
DB 756 GCGAGATTCCTCGGGATATACCCCATCAGTCACTAACGCCAGATCTCAGGAGAAATCC 815

QY 544 CTGGAGCTTGGCTGATCATCATGATGATGCCCTACGGGGAAACATCAGACAGCTGCTC 603
DB 816 CTGGTCTTCCCTGACCATCATATGATGATGCCCACTGGGAACATTTGGCAAGCGTCTC 875

QY 604 CGTTTCCCTATGTGAATCATTCACCAATCCGCTCAAGGGAGTCTCCGGCAGTATTGGGA 663
DB 876 CATTTCCCTATGTGAACCAATTCGCCAAATCCGCTCAAGGGAGTCTCTGGTAGCAATTGGGA 935

QY 664 ATGTGCTCGGAAAGAGAGAGTTGATGGCTGGAATAATGAAGCAGGACAAAGGGTGGCATTC 723
DB 936 ATGTGCTCGGAAAGAGAGAGTTGATGGCTGGAATAATGAAGCAGGACAAAGGGACCATTC 995

QY 724 CCAATGCTAATGGGACAGCAAGTGTCCCTCTGAAGGTCGGGAGCTACTGACATCGATG 783
DB 996 CCATGACGAATGGGACAGCAAGCAATGTCTCCCTCTGAGGGTCGGGGTGTGGTGATATTGATG 1055

QY 784 CATCTACTGAATACATGAGAGAGCGCTTACTGAAATGATGAACTCCGCCAGCTCTAT 843
DB 1056 CATCAACTGATTACATGAGAGATGCGCTTATTGAACGACGAAATCGACAGCCCTCTAT 1115

QY 844 CTAGAAAAGTCCCATTTGCTTCTCCAAAATAAAATCCCTACAGAAATGGTCATTGTTCTGC 903
DB 1116 CTAGGAAAGTTCACATCTCTCTCCAGGATAAATCCATACAGGATGGTCATTGTGCTGC 1175

QY 904 GGTGGTGTCTTAAAGCATCTTCTGACATCCGCTCTCAAAATCTGTGGGTAATGCAAT 963
DB 1176 GATTGATGTTCTTAAAGCATCTTCTGACATCCGCTATCAAAATCTGTGGGCAATGCAAT 1235

QY 964 ACCACTGTGGCTTTTATCTGTATATGAGATTGGTTGGTTGCTTTATCTGATGACTGG 1023
DB 1236 ACCCATATGGCTTCTATCTGTATATGAGATCTGGTTGCTCTTTCTGCGATATGG 1295

QY 1024 ATCAGTTCGCCGAAGTGGTTTCCAATCAACCGGAGACCTACCTTGATAGACTGGCTTTAA 1083
DB 1296 ATCAGTTCCTTAAGTGGTTTCCAATCAACCGGAGAGCTACCTTGATAGCTGGCATTA 1355

QY 1084 GGTATGACCGGAGAGGTGAACCGTCTCAGTTGGCTGCTGTTGACATATTTGTCAGTACAG 1143
DB 1356 GGTATGACCGGAGAGGTGAGCCATCTCAGTTGGCTGCTGTTGACATATTTGTCAGTACAG 1415

QY 1144 TCGACCCCTTGAAGGAGGCCACTATGCTCACTGCCAACTGTGCTATCTCATCTCTTGTCTG 1203
DB 1416 TCGACCCAAATGAAGGAGCCCTCTCTTGTCACTGCCAAATACCGTGTCTATCCATTTCTGCTG 1475

QY 1204 TTGATATCCCTGGACAAGGCTCTCTTGTCTATGTATCTGATGACGAGCTTCAATGCTGA 1263
DB 1476 TGGATTTACCTCTGGATAGGCTCTCTTGTCTATGTATCTGATGATGAGCTGCGATGCTGA 1535

QY 1264 CTTTGTGACGCAATGGCTGAGACTTTCAGAGTTTGTCTAGGAAATGGGTGATCCATTTGTGAAGA 1323
DB 1536 CATTTGATGCACTAGCTGAGACTTTCAGAGTTTGTCTAGGAAATGGGTGATCCATTTGTGAAGA 1595

QY 1324 AGTATGACATTCGAACCCAGAGCTCCGAGTTTCTCTTTTGCAGAAAAATGATATCCTCA 1383
DB 1596 AGTACAACTTGAACCTAGAGTCTCTGAATGGTACTTCTCCAGAAAAATGATATCCTCA 1655

QY 1384 AAGACAAATCCAGCCTTCAATTTGTTAAAGACCGCGGCCATGAAGAGAGAAATGAAG 1443
DB 1656 AGGACAAATGTCACCTTCAATTTGTTAAAGACCGCGGCCATGAAGAGAGAAATGAAG 1715

QY 1444 AATTTAAATTCAGGATTAATGCCCTAGTTTCTAAGGCATGAAAGTCCCGAGGAAAGGAT 1503
DB 1716 AATTTAAAGTTAGGTTAAATGGCTTGTGTTAAGGCACAGAAAGTCTCTGAGGAAAGGAT 1775

QY 1504 GGATCATGCAAGATGSCACCATGSCCAGGAAACAAATACAGGGATCATCTCTGGAATGA 1563
DB 1776 GGATCATGCAAGATGSCACCATGSCCAGGAAACAAATACCGGGACCATCTCTGGAATGA 1835

QY 1564 TTCAGGTTTTCTTGTGTCACAGTGGTGGCTTTGATACCTGAGGTAATGAGCTCCCCGTT 1623
DB 1836 TTCAGGTTTTCTTGTGTCACAGTGGTGGCTTTGATACCTGAGGCAATGAGTACCCGTT 1895

QY 1624 TAGTTTATGTGCTCTGTTGAAAGCGTCTGGGTTTCAGCACCAAGAGGCTGGTGCCA 1683
DB 1896 TGCTCTATGTTTCTCGTGAAGCGCTCTGGATTCAGCATCAAGAAAGCTGGTGCCA 1955

QY 1684 TGAATGCCCTGTCTGCTGCTCAGCTGCTTACTAATGCAATAATCATGTTGCAATCTTG 1743
DB 1956 TGAATGCTCTGTTGCTGCTGCTCAGCTGCTTACCATGGAATATCATGTTGAAATCTTG 2015

QY 1744 ATTGTGATCACTACATCAACAAACAGCAGGCTGTCCGAGAAGCTATGTCTTCTCTAATGG 1803
DB 2016 ATTGTGATCACTACATCAACAAACAGTAAAGCTCTCAGGGAAGCTATGTGCTTCTCTTATGG 2075

QY 1804 ATCCAAACCTAGGTCGCAAGTCTGTATGTGAGTCTCCCAAAAGGTTTGATGGGATGG 1863
DB 2076 ACCCTAAACCTAGGAAAGGAGTGTCTGCTACGTCCAGTTTCCCGAGAGATTCGATGGCATGG 2135

QY 1864 ATAGGAATGATCGATATGCAAAACAGGAACACTGTCTTTTTCATATATACTTGAAGGGCC 1923
DB 2136 ACAGGATGATCGATATGCAACAGGAACACCGTGTGTTTTCGATATTAATCTTGAGAGGTC 2195

QY 1924 TTGAAGGCATTCAGGACCAAGTTTATGTGGGAACCTGGTTGTGTTTTCACACAGAAACAGCTA 1983

Db 2196 TTGATGGCATCCAAAGCAGTTTATGTGCGAACTGGCTGTGTTTCAACGGAAGAGCTC 2255
QY 1984 TCTATGGTTATGAGCCGCCAAATTAAGGCGAAGAACGAGTTTCTTGGCATCACTATGTG 2043
Db 2256 TATATGTTTATGAGCCGCCAAATTAAGCAGAGAAGAGGTGTTTCTTGTGTCATCACTATGTG 2315
QY 2044 GGGGCAAGAGGAGGCAAGCACTCAAGAAAGAGGAGCTCAGATAGAAAAGTTCGAACA 2103
Db 2316 GCGTAGGAAGAGGCAAGCAATCAAGAA ---GGGCTCGGCAAGAAAGTTCGACA 2372
QY 2104 AGCATGTGACAGTTCTGTTCCAGTATTCAATCTCGAAGACATAGAGGAGGTGTTGAAG 2163
Db 2373 AGCATGTGACAGTTCTGTGCGCAGTATCAACCTTGAAGATAGAGGAGGAGTTGAAG 2432
QY 2164 GTGCTGGGTTTGATGATGAGAAATAGTCTCATGTCTCAATGAGCTTAGAGAAGAGAT 2223
Db 2433 GGGCTGGATTTGACGACGAGAAATCACTTCTTATGTCTCAATGAGCTTGGAGAAGAGAT 2492
QY 2224 TTGGCCAGTCAGCAGCATTTGTTGCCCTCCACTCTGATGGAATATGTTGTTCTCAGT 2283
Db 2493 TTGGCCAGTCCGAGCGTTTGTGCTCCACTCTGATGAGTATGTTGTTCTCAGT 2552
QY 2284 CTTCCACTCCAGATCTCTTTTGAAGAAGCTATCCATGTCTATAAGTTGTGGCTATGAG 2343
Db 2553 CCGCAACTCGGAGTCTCTTCTGAAAGAAGCTATCCATGTATAAGCTGTGGCTATGAG 2612
QY 2344 ACAAGTCTGAATGGGAACTGAGATGTTGGATCTATGATCTGATGATCTGACAGAAATATTC 2403
Db 2613 ACAAGACTGAATGGGAACTGAGATCGGTTGATCTACGTTCTGTGACAGAAGACATTC 2672
QY 2404 TTACTGGATTCAAGATGACGCAAGAGGCTGGGTTTCAGTCTATTTCATGATCCCAAGCTCGGT 2523
Db 2673 TCACCGGATCAAGATGACGCGGAGGCTGGCGGTGCGATCTACTGATGATCCCAAGCGGC 2732
QY 2464 CAGCTTTCAGGGAATGCCCCCATCAATCTTTTCAGATCGTCTGAAACCAAGTCTCGGT 2523
Db 2733 CAGCTTTCAGGGAATGCCCCCATCAATCTTTTCGACCGCTCTGAACCAAGTCTCGGT 2792
QY 2524 GGGCTCTGGTCTGTTGAAATCTTTTCAGCGGCAATGCGCTTATGATGATGATGATG 2583
Db 2793 GGGCTCTGGTCTGGGATGCTCTTCAGCGGCAATGCGCTTATGATGATGATGATG 2852
QY 2584 GAGGGCGCTCAAGTTCCTGGAGATTCGTTTACATCAACCAACCACTTACCACTAA 2643
Db 2853 GAGGGCGCTCAAGTTCCTGGAGATTCGTTTACATCAACCAACCACTTACCACTAA 2912
QY 2644 CCTCTCTCCGCTCTAGTCTATTGATATGCTGCTCTATCTGCTCACTGGAAGT 2703
Db 2913 CGTCCATCCGCTCTCTCATCTACTGATCTGCGCCGCTCTGCTGCTCAGCGAAGT 2972
QY 2704 TCATCATGCGAGATAGCACTTGGCAGTATCTGTTTCAATGCGCTCTCTCTTCAA 2763
Db 2973 TCATCATTCAGAGATCAGCACTTCCGAGATCTGTTTCAATCTCTCTCTCTCTCA 3032
QY 2764 TTTTGGCCACTGATCTCTGATGATGAGTGGTCTGCGATGAGGATGAGTGA 2823
Db 3033 TCTTCCGCAAGGATCTCTGGAGATGAGTGGGAGCGGGTGGGATCGAGAGTGGTGA 3092
QY 2824 GGAATGAACAGTTCTGGGTCATTGGAGGATCTCTGACATCTGTTTGGCTCTTTTCAG 2883
Db 3093 GGAACGAGAGTTCTGGGTGATCGGGGATCTCCGCGACCTCTTCGCGGTTCAGG 3152
QY 2884 GTCTTCTGAAGTCTTGGCGTATCGACCAACTCTGTCATCTCAAGGCTTAAG 2943
Db 3153 GCGTGTCTAAGGTCTGGCGGATCGACCAACTTCCCGTCACTCCAAAGGCTCG 3212
QY 2944 ACGAAGAGCGACTTGTGAGCTCTACATGTTCAAGTGGAGCGCTTCTCATCCCTC 3003
Db 3213 ACGAGACCGGACTTCCGGAGCTGATCATGTTCAAGTGGAGCGCTCTCTGATCCCG 3272
QY 3004 CGACGACCAATTTGATCATTAACATGTTGGTGTGCTGTCGCACTCTCTACGCCATCA 3063

Db 3273 CCACCACCATCTGATCATCAACCTGGTGGCGTCTGCGCGCATCTCTACGCCATCA 3332
QY 3064 ACAGTGGTTACCAATCATGGGGCCGCTCTTTGGGAAGCTCTTTTGGCTTCTGGTGA 3123
Db 3333 ACAGCGGATACAGTCTGGGGCCGCTCTTCGGCAAGCTCTTCTCGCCTTCTGGTCA 3392
QY 3124 TTGTTCACTTATACCATTCCTCAAGGCTCTTATGGGAGGCAAAACCGCACACCGACA 3183
Db 3393 TGTCCACCTGTACCCGTTCTCAAGGGCTCTATGGGAGGCAAAACCGCACACCGACA 3452
QY 3184 TTGTCACTCTGGGGCTCTCTCGCTTCTATCTTCTCTTCTGCTGCTGCTGCTGCTG 3243
Db 3453 TCGTCTGCTCTGGGGCATCTCTGCTGGCTCCATCTTCTCTGCTGCTGCTGCTGCTG 3512
QY 3244 ATCCATCTACTACCGTCTCGCTGGCCCAAAATCAAACTGTGGCATCAACTGTAGG 3303
Db 3513 ACCCTTCAACCCCGCTCACTGGCCCGATACCCAGACGTTGGCATCAACTGTAGG 3572
QY 3304 AAAGTGGGAGTTT 3316
Db 3573 GAAGTGAAGGTT 3585

RESULT 6
ADA69499
ID ADA69499 standard; DNA; 3222 BP.
XX ADA69499;
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 2822.
XX Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
OS Oryza sativa.
XX
XX WO2003000898-A1.
XX
XX 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 6; SEQ ID NO 2822; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.
XX Sequence 3222 BP; 805 A; 711 C; 805 G; 901 T; 0 U; 0 Other;

Db 2129 GCTTAGAGAAAAGATTGGTCAATCTAGTGTGTTTTTGTAGCCTCCACTCTGATGGAATATG 2188
QY 2269 GTGGTGTCTCAGTCTCCTCAGATCTCAGAAATCTCTTTTGAAGAAGCTATCCATGTCAFAA 2328
Db 2189 GTGGTGTCTCCTCAATCCGCAACTCCAGAAATCTCTTTCTGAAGAGGCCATACATGTTATCA 2248
QY 2329 GTTGTGGCTATGAGGACAGTCTGAATGGGAACATGAGATTGGTTGGATCTATGATCTG 2388
Db 2249 GCTGTGGCTATGAGACAAAGTGAATGGGAACATGAGATTGGATGGAATCTATGTTCTG 2308
QY 2389 TCACAGAAGATATCTTACTGGATTCAAGATGACGCAAGAGCTGGCGTTTCAGTCTATT 2448
Db 2309 TTACAGAAGATATCTTACTGGATTCAAGATGACGCAAGAGCTGGCGTTTCAGTCTATT 2368
QY 2449 GCATGCCAAGCGCCAGCTTTCAAGGATCTGCCCCCAATCAATCTTTTCAGATCTGCTGA 2508
Db 2369 GCATGCCAAGCGCCAGCAATTTCAAGGGGTCTGCCCAATTAATCTTTTCGGATCTGCTGA 2428
QY 2509 ACCAAGTGTCTGGGTGGCTCTCGGTTCTGTTGAAATCTTTTTCAGCGGCAATTTGCCCT 2568
Db 2429 ACCAGGTGCTTCGGTGGGTCTTGGTTCTGTTGAAATCTTTTTCAGCGGCAATTTGCCCA 2488
QY 2569 TATGTTATGGCTACGAGGCGGCTCAAGTTCCTGGAGATTCGCTTACATCAACACCA 2628
Db 2489 TATGTTATGGGTACGAGGAGGCTTAACTTCTGGAGAGATTGCTTACATCAACACCA 2548
QY 2629 CCATTTACCACTAACCTCTCTCCGCTCTAGTCTATTTGATATTTGCTGCTATCTGTC 2688
Db 2549 CTAATTTACCACTACATGATTCACCTCTCTTGTACTGTCATATTTGCCAGCTATCTGTC 2608
QY 2689 TGCTCACTGGAAGTTTCATGCGCAGAGATAGCAATTTGCGGATCTCTGTTTCATG 2748
Db 2609 TTCTCACTGGGAGTTTCATACACAGAGATAGCAATTTGCGGATCTCTGTTTCATG 2668
QY 2749 CGCTCTCTCTTCAATTTTCGCACTGCTATCTTGGAGATGAGTGGAGTGTGTTGCA 2808
Db 2669 CTCTCTCTCTGTCATTTTGTCTACTGTTATCTTGGAGATGAGTGGAGTGTGTTGCA 2728
QY 2809 TTGAGAGTGTGGAGGAATGAACAGTTCTGGGTCAATGGAGGTATCTCTGACATCTGT 2868
Db 2729 TCGATGAATGGTGGAGGAATGAGCAATTTTGGGTATTTGGAGGTATTTCCGCTCATCTAT 2788
QY 2869 TTGCGGTCTTTCAGGCTCTCTGAGGTCTTGGAGGTATCGACACCACTTTCATGTTCA 2928
Db 2789 TTGCGGTCTTCAGGCTCTCTGAGGTCTTGGAGGTCTTGGAGGTATGATACCACTTTCATGTTCA 2848
QY 2929 CTTCAAGCTTAATGACGAAGAGCGACTTTGCTGAGCTTACATGTTTCAAGTGGACGA 2988
Db 2849 CTTCAAGGCTCTGATGAAGAGCGGATTTGCTGAGCTTACATGTTTCAAGTGGACAA 2908
QY 2989 CGCTTCTATCCTCTCGGACGACATTTGATCATTTAAAGTGTGGTGTGCTGTTGCA 3048
Db 2909 CCGTTCTGATCCACCGACCACTATTTTATCATCAACCTGCTGGGTGTGTTGCTGTTA 2968
QY 3049 CTTCTACGCCATCAACAGTGTGTACCAATCATGGGGCGGCTCTTTGGGAAGCTTCTCT 3108
Db 2969 TCTCTACGCTATCAACAGTGTGTACCAATCATGGGGCGGCTCTTTGGGAAGCTTCTCT 3028
QY 3109 TTGCGCTTCTGGGTGATTTGTTCACTTATACCCATTTCTCAAGGCTCTTATGGGACGCAAA 3168
Db 3029 TTGCGCTTCTGGGTGATTTGTTCACTTATACCCATTTCTCAAGGCTCTTATGGGACGCA 3088
QY 3169 ACCGACACCGAGGATTTGATGCTGTGGGTGTCTCTCTGCTGTCTATCTCTCTCTTCTG 3228
Db 3089 ACCGTACACCGACCAATTTGTTGTTTGGGCCATCTCTCTGATCGATCTTCTCTCTG 3148
QY 3229 TGTGGGTCTGTGTTGATTCATTCATTCACCTGCTGCTGCGCCCAATTCACCACTGTG 3288
Db 3149 TGTGGGTCTGTGATTCATTCATTCACCTGCTGCTGCGCCCAATTCACCACTGTG 3208
QY 3289 GCATCAACTGCTAG 3302
Db 3209 GTATCAACTGCTAG 3222

RESULT 7
AAZ99533
ID AAZ99533 standard; DNA; 3704 BP.
XX
AC AAZ99533;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
FH Key Location/Qualifiers
FT CDS 272..3496
FT /*tag= a
FT /product= "cellulose synthase"
FT /transl_except= (pos: 1546, aa: Xaa)
FT /note= "no termination codon given; Xaa is an unspecified
XX amino acid"
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
PR 17-AUG-1998; 98US-0096822P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
DR P-PSDB; AAY84121.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
XX
PS Claim 1; Page 191-196; 119pp; English.
XX
CC The present sequence encodes a maize cellulose synthase polypeptide. The
CC cellulose synthase can be used for the improvement of stalk quality for
CC improved stand or silage. It also provides an increased concentration of
CC cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants and
CC seeds expressing the cellulose synthase. The polynucleotide is used for
CC modulating, preferably increasing, the level of the synthase in a plant
CC cell. The plants are preferably monocots. The polynucleotide is also used
CC as a probe or primer in the detection quantitation or isolation of gene
CC transcripts. The probes are useful in detecting deficiencies in the level
CC of mRNA in screenings for desired transgenic plant, for detecting or
CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX
SQ Sequence 3704 BP; 909 A; 838 C; 962 G; 994 T; 0 U; 1 Other;
Query Match 64.1%; Score 2325.2; DB 3; Length 3704;
Best Local Similarity 82.9%; Pred. No. 0;
Matches 2686; Conservative 0; Mismatches 544; Indels 9; Gaps 3;

Db 2426 GGCAGTCCAGTGTGTTTGTAGCCTCTACTCTGATGGAAATATGGTGGTTCCTCCAAATCT 2485
Qy 2286 TCCACTCCAGAAATCTCTTTTGAAGAAGCTATCCATGTCAATAGTTGGTGTATGAGGAC 2345
Db 2486 GCAACTCCAGAGTCTCTTCTGAAGAAGCTATTCATGTATCATCAGCTGTGGCTATGAGGAC 2545
Qy 2346 AAGTCTGAATGGGAACTGAGATTGGTTGGATCTATGGATCTGTCTACAGAGATATTTCT 2405
Db 2546 AAAACTGACTGGGAACAGATGGGTGGATCTATGGTCTGTGTACAGAGACATTTCTC 2605
Qy 2406 ACTGATTCAGATGACGCAAGAGCGTGGCTTCAGTCTATTGCTATGCCAACGCGCCA 2465
Db 2606 ACCGATTCAGATGATCTGCGAGCTGGCAATCAATCTACTGATGCTTAAGCAGCA 2665
Qy 2466 GCTTTCAAGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTGTGGGTGG 2525
Db 2666 GCTTTCAAGGATCTGCTCTTATCAACTTTTCGGATCGTTTGAATCAAGTGTCTGGTGG 2725
Qy 2526 GCTCTGGTCTGTGAAATCTTTTCAGCGGCAATGCGCCCTTATGGTATGGCTACGGA 2585
Db 2726 GCTCTGTGTTCAATGAAATCTTTTCAGAGCGCAATGTCATATGGTATGGCTATGGA 2785
Qy 2586 GGGCGCTCAAGTCTCTGAGAGATTGCTTACATCAACACCACTTTACCCACTAAC 2645
Db 2786 GCGCGCTTAAATCTCTGGAGAGATTGCTTATATCAACACAAATTTATCCATCA 2845
Qy 2646 TCTCTCCGCTCTAGTCTATTGTATATGCTGTCTATCTGTCTCACTGGAAGTTC 2705
Db 2846 TCAATCCCGCTCTCTCTGACTGATATTTGCGAGAGTTTGTCTCACTGGGAAGTTC 2905
Qy 2706 ATCATGCCAGATTAGCAACTTGGCCAGTATCTGGTTCATTGCGCTCTTCTTCAATT 2765
Db 2906 ATCATCCCAAGATTAGTAAACCTAGAGAGTGTGGTTTATATGCTCTTTATCAATC 2965
Qy 2766 TTGCGCACTGTPATCTCTGAGATGAGTGGAGTGTGGATTCACAGATGGTGGAGG 2825
Db 2966 TTGCGCACTGTPATCTCTGAGATGAGTGGAGTGTGGATTCACAGATGGTGGAGG 3025
Qy 2826 AATGAACAGTCTGGGTCTATGAGGATFCTCTGCACAFCTGTTGGCTCTTTCAGGT 2885
Db 3026 AACGAGAGTCTGGGTCTATGTTGTTGTTCTGGCAATTTATTTGGCTCTTTCAGGT 3085
Qy 2886 CTTCTGAGGTCTTGGCGGTATGACACCAACTTCACTGTCACTCAAGCTATATGAC 2945
Db 3086 CTCCTGAAGGTCTTGGTATGACACAGAGCTTCACTGTCACTCTAAGGCCACTGAC 3145
Qy 2946 GAAGAAGCGACTTTGCTGAGCTTACATGTTCAAGTGACGACGCTTCTCATCCCTCG 3005
Db 3146 GAAGAAGGTGATTTGCGAGCTTACATGTTCAAGTGACGACGCTTCTGATCCACCA 3205
Qy 3006 ACGACCATTTGATCAATTAACATGTTGGTGTGGTGTGGACCTTCTTACGCCATCAAC 3065
Db 3206 ACCACTATTTGATCATCAACCTGTCGGGTGTGGTGTGGATTTCTTACGCCATCAAT 3265
Qy 3066 AGTGTTCACCATCATGGGGCGCTCTTTGGAGCTCTTCTTGGCTCTTGGGTGAT 3125
Db 3266 ACGCGTACCATGTCATGGGACCTCTTTTCGGGAAGCTCTTCTTGGGTCTTGGGTGAT 3325
Qy 3126 GTTCACTTATACCATCTCTCAAGGTCTTATGGGAGCGCAAAACCGCACACCGAGAT 3185
Db 3326 GTCCACTGTACCCCTCTCTCAAGGCTCTATGGGAGAGAGACGACGACGACCAT 3385
Qy 3186 GTCATGCTCTGGGCTCTCTCTCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3245
Db 3386 GTCTGTCTGGGCTATCT 3445
Qy 3246 CATTCACCTACCGCT 3304
Db 3446 CCATTCACCAACCGGCT 3504

RESULT 8

AAZ58271
ID AAZ58271 standard; cDNA; 1734 BP.
XX
AC AAZ58271;
XX
DT 08-MAY-2000 (first entry)
XX
DE Wheat cellulose synthase cDNA contig.
XX
KW Wheat; cellulose synthase; transgenic plant; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 2..1522
FT /*tag= a
FT /partial
XX
PN WO200004166-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015871.
XX
PR 14-JUL-1998; 98US-0092844P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;
PI Rafalski JA, Thorpe CJ;
XX
DR WPI: 2000-182431/16.
DR P-PSDB; AAY58840.
XX
PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
probes for isolating cDNAs and genes encoding homologous proteins, for
producing transgenic plants.
XX
PS Claim 3; Page 66-67; 93pp; English.
XX
CC The present sequence is that of a contig of cDNA clones and encodes a
portion (see AAY58840) of wheat cellulose synthase (CS). The cDNA clones
were isolated from wheat 7-day seedling leaf and 7-day etiolated seedling
root tissue cDNA libraries on the basis of homology to Arabidopsis and
cotton CS sequences. The invention relates to isolated nucleic acid
fragments encoding plant CS and to CS polypeptides. It also relates to
the construction of a chimeric gene encoding all or a portion of the CS,
in sense or antisense orientation, where expression of the gene results
in altered levels of the CS in transformed host cells. The host cells can
be used to screen compounds for their ability to inhibit CS activity. CS
nucleic acids are also useful for producing transgenic plants having
altered levels of CS, and hence altered levels of fibre. CS may also
serve as a target for the development of novel herbicides
XX
SQ Sequence 1734 BP; 441 A; 369 C; 420 G; 502 T; 0 U; 2 Other;
Query Match 46.4%; Score 1681.6; DB 3; Length 1734;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1696; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
Qy 1782 GAAGCTATGCTGCTTCTCTATGATCCAAACCTAGGTCGCAAGTCTGTTATGTCAGTTC 1841
Db 2 GAAGCTATGCTGCTTCTCTATGATCCAAACCTAGGTCGCAAGTCTGTTATGTCAGTTC 61
Qy 1842 CCACAAAGTTTGTATGGGATTCAGGATGATAGGATGATGCAACAGGAACTGTCTTT 1901
Db 62 CCACAAAGTTTGTATGTTATTCATAGGAATGATCGATATGCAACAGGAACTGTCTTT 121
Qy 1902 TTTGATATTAACCTTGAAGGCGCTTCAAGGCATTCAGGCACTTATGTCGGAACGTG 1961
Db 122 TTTGATATTAACCTTGAAGGCGCTTCAAGGCATTCAGGCACTTATGTCGGAACGTG 181
Qy 1962 TGTGTTTTTCAACAGAACAGCTATCTATGTTATGAGCCCCCAATTAAAGCGGAAGCCA 2021

[illegible]

RESULT 9

RESULI 3
AAC49550

AAC49330
ID AAC49550 standard: DNA: 3198 BP.

AAC AAC49550:

18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana DNA fragment SEO ID NO: 61572.

AA Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

06-SEP-2000.

25-FEB-2000: 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.
 EE 01 APR 1999; 99US-0127103P.
 EE 01 APR 1999; 99US-0127103P.

PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.

PK 08-APR-1999; 99US-0128714P
PK 08-APR-1999; 99US-0128714P

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.

PK 30-APR-1999; 99US-0132048F;
PR 30-APR-1999; 99US-0132407P

PR 04-MAY-1999: 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134769P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0132487P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159323P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.

Query Match 46.3%; Score 1677.6; DB 3; Length 3198;
Best Local Similarity 72.0%; Pred. No. 0;
Matches 2330; Conservative 0; Mismatches 834; Indels 72; Gaps 8;

PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

77 CGCCCTGAAGTCGGGAGGACGCGGGCCGGGACGTGTGCCAGATCTGCCGACGCGCT 136
21 CGCGGAAAGCCGATGAAGAACATTTTCGCGACTTGCAGATCTGTAGTGAATGT 80
137 GGGCACACAGTGTGACCGGACGCTCTTCAACGCTCGGACGCTCTGCCGCTTCCCGCTG 196
81 TGGCAAGACTGTGTGATGGAGATCGTTTGTGGCTTGTGATATTTGTTCAITTCGCCAGTTG 140
197 CGGCCCTGTACGAGCACAGCGCAAGGAGGACCGACCGCTCGCTCGAGTCCAGAC 256
141 TCGGCCCTGTACGAGTATGAGAGGAAGATGGGAATCAATCTTGTCTCAGTCAAAAC 200
257 CAAGTCAAGCCACAGAGGAGCCGAGCGATCCGGGGAGGAAGCGGACGACACTGA 316
201 CAGATCAAGAGGCTCAAGGTAGTCTGTATTTCTGTGTGATTAAGACGAGGTGCTT 260
317 TGGCGATGATGGTA---GTGACTTCAACTACCTTGCATCTTGCACTGAGGACCAAGACA 373
261 AGCTGATGAAGTACTGTGAGTCAACTACCTT-----CAGAAGGA 302
374 GAAGATTGCTCAGAGGATGCGCAGTCCGCGATGAACACCGGGGAGTGGCAATGTGG 433
303 GAAATTTACAGCGGATGCTGTGTGGCATCTTACTCGTGGGAGGAGGAAATGGG 362
434 CCACCCCAAGTATCAGAGTGGCGAGATCGGCTCTCCAAAGTATCACAGTGGAGATCCC 493
363 GBAACCCAGTATGATGAAGG-----TCTCTCAATCATCTTCTCTGCTCAC 413
494 TAGGGATAGTCCCTTTCAGTCAACCAAGCCAGATGTCAAGGAATTCCTTGAGCTTC 553
414 GAGCAGACAAG-----ATACTTCAGGAGAGTTTCTGCTGCCTC 452
554 GCCTGATCATCATGATGTCCCTACGGGAAATCATCAGAGAGTCTCGTTTCCT- 612
453 ACCTGAAGGCTCTCTGTATCTTCTATATGCTGTGGGAAAGCGCTTCCCTATTATC 512
613 --ATGTGAATCATTCACCAATCGTCAAGGAGTCTCCGGCAGTATGGGAATGTTC 670
513 AGATGTCATCAATCAACCAATAGAAGATTGTGGATCTTGGACTCGGGAATGAGC 572
671 CTGAAAGAGAGGTTGATGGCTGGAATAATGAAGCAGCAAGGGTCCGATTCCTCATGAC 730
573 TTGGAAGGAGAGGTTGATGGCTGGAATAATGAAGCAGCAAGGAACTACTGGTCTGTGAC 632
731 TAATGGACAAGCATCTCCCTCTGAAGGTCCGGGAGTACTGACATCGATGATCTAC 790
633 CAGCAG-----GCTGCTTCTGAAGAGGTGGAG---TAGATATTGATCCAGCAC 680
791 TGAATCAACATGAAGACGCTTTCATGAATGATGAATACTCGCAGGCTCTATCAGAAA 850
681 AGATATCTAGCAGATGAGGCTCTGCTGAATGACGAAGCGAGGCGAGCTCTGTCAAGGAA 740
851 AGTCCCATGCTTCTCCAAATAAATCCCTACAGAAATGGTATTGTTCTGCGGTTGGT 910

Db 741 AGTTTCAATTCCTTCATCAGGATCAATCCTTACAGAAATGTTATTATGTCGCGCTGT 800
Qy 911 TGTCTAAGCATCTTCTCTGCACTACCGTCTCAAAATCTGTGGTGAATGATACCCACT 970
Db 801 TATCCTTTGTCTCTTCTTGTGATTAACCGTATAACAAACCCAGTGCCTTTGTCTCT 860
Qy 971 GTGGCTTTTATCTGTTATATGAGATTTGTTGTTTCTTCTGGAATCTGGAATCAGTT 1030
Db 861 ATGGCTGGTCTCTGTGATATGTGAGATCTGGTTTTCCTTATCTCTGGAATTTTGGATCAGTT 920
Qy 1031 CCCGAAGTGGTTTCCAAATCAACCCGGGAGACCTACTCTGTATAGACTGGCTTTAAGATGTA 1090
Db 921 TCCCAAGTGGTTTCTCTGTGAACCGTGAAACCTACTCGACAGGCTTCTCTTAAAGATGA 980
Qy 1091 CCGAAGGTGAACCGTCTCAGTTGGCTGCTGTGATGACATATTTGTCTAGTACAGTCGACCC 1150
Db 981 TCGTGAAGGTGAGCCATCACAGTTAGTCTGTGTTGACATTTTCGTGAGTACTGTGTACCC 1040
Qy 1151 CTTGAAGGAGCCACTATCGTCACTGCCAACACTGTCTATCCATTCCTTGTCTGTGATTA 1210
Db 1041 CTTGAAGGAGCCACCCCTCTGTGACGCCAACACAGTCTCTCTATCTGCTGTGTGACTA 1100
Qy 1211 TCCCGTGGACAAGTCTCTTCTGTATGATCTGATGACGAGGCTTCAATGCTGACTTTTGA 1270
Db 1101 CCCAGTTGACAAGGTGCTGTTATGTTCTGATGATGGTCTGCTATGTTATCATTTGA 1160
Qy 1271 CGCATTTGGCTGAGACTTTCAGAGTTTGTAGGAATGGTACCATTTTGTGAGAGTATGA 1330
Db 1161 ATCATTTCGAGAAACATCAGAGTTTGTCTGTAATGGGTACCATTTTGCAGAAATATAG 1220
Qy 1331 CATTGAACCCAGAGCTCCCGAGTTTACTTTTTCGAGAAAATGATTACCTGAAAGACAA 1390
Db 1221 CATAGAGCTCTGTGCACAGAAATGGTACTTTTGTGCGAATAATAGATTACTTTGAAGATAA 1280
Qy 1391 AGTCCAGCTTCAATTTGTTAAAGACCGCGGCGCATGAAGAGAAATATGAAGAAATTA 1450
Db 1281 AGTTCAGACATCATTTGTCAAAGATCGTAGAGCTATGAAGAGGAAATATGAGAAATTA 1340
Qy 1451 AATCAGATAAATGCCCCCTAGTTTCTAGGCATTTGAAAGTCCCGAGGAGGATGATCAT 1510
Db 1341 AATCCGAATCAATGCACTTTGTTCCAAAGCCCTTAAATGTCTTGAAGAGGTTGGTTAT 1400
Qy 1511 GCAAGATGGCACACCATGGCCAGAAACAATACAGGGAATCATCTCGAATGATTCAGGT 1570
Db 1401 GCAAGATGGCACACCGTGGCTGGAATATACAGGGAATCATCCAGGAATGATCCAGT 1460
Qy 1571 TTTCTTTGGTCAAGTGGTGGCTTGTATATGAGGGTAAATGAGCTCCCCGTTTGTGTTA 1630
Db 1461 CTTCTTAGGGCAAAATGGTGGACTTGTATGACAGAGGCAATGAGCTCCCGCTTTGGTATA 1520
Qy 1631 TGTGCTCTCGTGAAGAGCTCTGGTTCAGGACCCAGAGGCTGGTCCCATGATGC 1690
Db 1521 TGTCTTCGAGAAAGCGACAGGATTCAGACCCACAAAGGCTCGTCTATGAATGC 1580
Qy 1691 CTTGTTCTGCTCTCAGCTCTCTTAAATGACAAATACATGTTGAATCTTGTATTTGTA 1750
Db 1581 ACTGGTGAAGTTTTCAGAGTCTTCAACATGGACCTTTTCATCTTGAATCTTGTATGTA 1640
Qy 1751 TCACTACATCAACAAAGAGAGGCTGTCCGAGAGCTATGCTTCTCTATGATGATCCAAA 1810
Db 1641 TCATTACATAAATAACAGCAAGCCCTTAAGAGAGCAATGTGCTTCTGATGGACCCAAA 1700
Qy 1811 CCTAGGTCCCAAGTCTGTTATGTGAGTTTCCCAAGAGTTTGTATGGGATTCGATAGAA 1870
Db 1701 CTTCCGGAAGCAAGTTTGTATGTTCTAGTTCCCAAGAAATTTGATGCTATCGATAGAA 1760
Qy 1871 TGAATGATGCAAAACAGGAACACTGCTGTTTTTTTGTATTAATTAAGTGGGCTTTGACGG 1930
Db 1761 CGATAGATATGCTAATCGTAATACCGGTGTTCTTTGATATTAATTAAGTGGGCTTTGATGG 1820
Qy 1931 CATTCAAGGACAGTTTATGTGGGAATGTTGTTGTTTCAACAGAACACTATCTATGG 1990
Db 1821 GATTCAGGACCTGTATATGTCGGAATGTTGTTTTCACAGAACAGCATTTATACGG 1880

QY 2058 GCAACCAAGTCAAAGAAAGAGCTCAGATAGAAAAGTCAAGCAAGCATGTGACAGT 2117
Db 2170 AAGAAATCCAAAGCTAAGAAAGAGTGGACAAAAAGAAATCAGGAGGATACTGACTCA 2229
QY 2118 TCTGTTCCAGTATTCATCTCAAGACATAGAGGAGGGTGTGAAGGTGCTGGTGTTCAT 2177
Db 2230 ACTGTTCTGTATTCATCTCAAGACATAGAGGAGGGTGTGAAGGTGCTGGTGTTCAT 2289
QY 2178 GATGAGAAATAGTTCATGCTCAATGCTTAGAGAGAGATTTGGCCAGTCAGCA 2237
Db 2290 GATGAAAAGGCGCTCTTAATGTCGAAATGAGCCTGGAGAGGATTTGGACAGTCTGCT 2349
QY 2238 GCATTTGTTGCTCCACTCTGATGAGATATGCTGTTGCTCCTCAGTCTCCACTCCAGAA 2297
Db 2350 GTTTTGTGTTCTACCTTAATGAGAAATGGTGGTTCCTCTTCCAGCAACTCCAGAA 2409
QY 2298 TCTCTTTTGAAGAAGCTATCCATGCTCAATGATGTTGGCTATGAGGACAAGTCTGAATGG 2357
Db 2410 AACTTTCTCAAGAGGCTATCCATGCTCATGTTGTTGTTATGAGGATAAGTCAGATTGG 2469
QY 2358 GGAACGTGAGATTGGTTGGATCTATGATCTGTCACAGAAAGATTTCTACTGATTCAG 2417
Db 2470 GGAATGGAGATTGGATGGATCTATGTTCTGTGACAGAAAGATATCTGACTGGGTTCAA 2529
QY 2418 ATGCAGGCAAGAGGCTGGGTTGAGTCTATTCATGATGCCCAAGGCCAGCTTTCAGGCA 2477
Db 2530 ATGCATGCCCCGTTGGATGGGATCCATTTACTGATGCCCTTAAGCTTTCAGGGT 2589
QY 2478 TCTGCCCCCATCAATCTTTTCAGATCTCTGACCAAGTCTGCGTGGGCTCTCGGTTCT 2537
Db 2590 TCTGCTCTCTCAATCTTTTCAGATCTCTGACCAAGTCTGAGGTGGCTTTAGTTTCA 2649
QY 2538 GTTGAATCTTTTCAGCCGGATTCGCCCTTANGTATGGCTACGGAGGGCGCCCTCAG 2597
Db 2650 GTTGAGATCTCTCAGTCGGCATTCCTATATGTTATGTTTACAAATGGAGGCTAAAA 2709
QY 2598 TTCTGGAGAGATTCCGTTACATCAACACACCATTTACCCATACCTCTCTCCGCTT 2657
Db 2710 TTCTGTGAGAGGTTTGGTATGGAACACACCATCTACCTATACCTTCCATTCCTCTT 2769
QY 2658 CTAGTCTATTGCTATTTGCTGCTATCTGCTCACTGGAAGTTCATCATGCCAGAG 2717
Db 2770 CTCATGATTTGATATGCTAGCCGTTGCTCTTCCACCAACAGATTTATTTCTCTCAG 2829
QY 2718 ATTAGCAACTTGCCAGTATCTGTTTCATTCGCTCTTCTTTCAATTTTGGCACTGGT 2777
Db 2830 ATTAGTAACATTGCAAGTATATGGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2889
QY 2778 ATCCTTGAGATGAGTGGAGTGTGTGGCATTGACGAGTGGTGGAGGATGAACAGTTTC 2837
Db 2890 ATACTAGAATAGGTGGAGTGGCTAGGCATAGAGCAATGGTGGAGAAACAGAGCGATT 2949
QY 2838 TGGGTCAATTGGAGGTATCTCTGCACATCTGTTTGGCGTCTTTTCAGGGTCTTCTGAAGG 2897
Db 2950 TGGGTCAATTGGTGGATATCCGCTCATTTATTCGCTGTGTTTCAAGGTATCCTCAAAGTC 3009
QY 2898 CTTGCGGTATGACACCAACTTTCATGTCCTCAAGGCTAATGACGAGAAAGGCGAC 2957
Db 3010 CTTGCGGTATGACACCAACTTTCATGTCCTCAAGGCTTCAAGGCTTCAAGTGAAGCGGAG 3069
QY 2958 TTTGCTGAGCTTACATGTTCAAGTGGACGACGCTTCTATCCTCCGACGACATTTTG 3017
Db 3070 TTTGCTGAGCTTACTTGTTCAAATGGACAAACATCTCTGATTCGCCAACGACGCTGCTC 3129
QY 3018 ATCAATTAACATGTTGTTGCTGCTGCGACCTCTTACGCCATCAACAGTGGTTACCA 3077
Db 3130 ATTGTAACATTAGTGGAGTGTGTCAGGAGTCTCTTATGCTATCAACAGTGGATACCA 3189
QY 3078 TCATGGGGCGCTCTTTGGAAAGCTCTCTTTGCTCTTGGGTGATGTTGCTATTATAC 3137
Db 3190 TCATGGGGACCACTCTTTGGTAAGTTGTTCTTTGCTCTTGGGTGATGTTGCTATTGAC 3249
QY 3138 CCATTCCTCAAGGTTCTTATGGGCGAGCAAAACCGCACCGGAGGATTGCTATGCTGCTG 3197

Db 3250 CTTTCTCAAGGTTTGTGATGGTGCAGAGAACCGGACTCTTACCATTGTTGTGCTGG 3309
QY 3198 GCTGTCCTCTCGCTTCTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3257
Db 3310 TCTGTTCTCTTGGCTTCTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3369
QY 3258 CTTCTCGTGGCCCAAAATATCCAACTGTGGCATCAACTCACTGCTAGAA 3305
Db 3370 CGAGTCACTGGCCCGGACATCTTGAATGTGAATCAACTGTGAGAA 3417

RESULT 11
AAZ99500
ID AAZ99500 standard; DNA; 3725 BP.
XX
AC AAZ99500;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 179..3398
FT /*tag= a
FT /product= "cellulose synthase"
FT /note= "no termination codon given"
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
PR 17-AUG-1998; 98US-0096822P.
XX
(PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
DR P-PSDB; AAY84110.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
XX
PS Claim 1; Page 105-110; 119pp; English.
XX
CC The present sequence encodes a maize cellulose synthase polypeptide. The
CC cellulose synthase can be used for the improvement of stalk quality for
CC improved stand or silage. It also provides an increased concentration of
CC cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants and
CC seeds expressing the cellulose synthase. The polynucleotide is used for
CC modulating, preferably increasing, the level of the synthase in a plant
CC cell. The plants are preferably monocots. The polynucleotide is also used
CC as a probe or primer in the detection quantitation or isolation of gene
CC transcripts. The probes are useful in detecting deficiencies in the level
CC of mRNA in screenings for desired transgenic plant, for detecting
CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of

CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX
SQ Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 U; 0 Other;

Query Match 38.8%; Score 1408.4; DB 3; Length 3725;
Best Local Similarity 66.8%; Pred. No. 0;
Matches 2147; Conservative 0; Mismatches 966; Indels 99; Gaps 6;

QY	53	CCGCGCATGACGACGCGGAGCGGAGCCCTGAAGTCCGGGAGGACACGGGGCCGGGAGCGT	112
DB	238	CCGCCACACGCGGACGCGCTGTCCCGGTGAAGCCACGAGAGTGCAGATGGGCGAGGT	297
QY	113	GTGCCAGATCTGCGCCGACGCGCTGGGCACACACCTTGGACGCGCACGCTTTCACCGCTG	172
DB	298	CTGCCAGATTTGTGGCGACACTGTTGGGTTTCAGCCACTGGTATGTCCTTTGTGCTG	357
QY	173	CGAGCTGCGCGCTTCCCGTCTGCGGCTCTGCGGCTCTGACGACGCGCAAGAGGGGAC	232
DB	358	CAATGAGTGTGCTTCCCTGTCTGCGGCTTGTCTATGAGTACGAGCGCAAGGAGGAA	417
QY	233	CCAGGCTGCTCCAGTGCAGACCAAGTACAAGGCCACAGAGGGAGCCGACGATCCG	292
DB	418	CCAATGCTGCCCTCAGTGCAAGACTAGATACAGAGACAGAAAGTACCCCTCGAGTTCA	477
QY	293	CGGGGAGGAGGCGACGACACTGATGCCGATGATGGTAGTGAATCAACTACCCCTGCATC	352
DB	478	TGGTGATGTAGGAGGAGAGATGTTGATGACCTGGACATGAATCAACTATAGCAAGG	537
QY	353	TGGCACTGAGGACGAGAGCAAGATGCTGACAGGATGCGAGCTGGCGCATGAACAC	412
DB	538	CAATGGGAAGGGCCGAGATGGCAGCTT-----CAAGGAGAT	574
QY	413	CGGGGCGAGTGCATGTTGSCCACCACCAAGTATGACAGTGGCGAGATCGSCCTCCAA	472
DB	575	GACCTGATCTGCTTCACTGCTCGCCATGACCCACCATCGATTCACAGCCTTACA	634
QY	473	GTATGACGTGGAGAGATCCCTAGGGGATAGTCCCTTTCAGTCCCAACAGCCACGATGTC	532
DB	635	AGTGGACA-----ACAGATATC	651
QY	533	AGGAAATCCCTGGAGCTTCCCTGATCATCATGATGCTCCCTACGGGGAAACATCAG	592
DB	652	TGGAGAGATCCCTGATCATCCCTGACCGCTCATTTATCCGAGTCCAAACATCGAGTA	711
QY	593	CAGAGCTGCTCCGTTTCCCTATGGAATCATTCACCAATCCGTCAGGGAGTT---CTC	649
DB	712	TGTTGATCCAGCGTTCAGTTCTGTGAGGATTTGTGACCCCTCGAAGGACTTGAATTC	771
QY	650	CGGCGATTTGGGAATTTGCTGGAAAGAGAGATTGATGGCTTGGAAATGAAGCAGGA	709
DB	772	CTATGGGCTTAATAGTTGCTGACTGGAAGGAAAGATTGAGAGCTGGAGGTTAAACAGGA	831
QY	710	CAAGGTTGCGATTCCTATGACTAATGGGACAGCAATGCTCCCTCTGAAGTCCGGCAGC	769
DB	832	CAAAAATATGTTGAAGTGACTAA-----TAAATATCCAGAGGC	870
QY	770	TACTGACATCGATGCTACTACTGAATACAAACATGGAAGACGCTTTACTGAATGATGAAC	829
DB	871	TAGAGGAGACATGGAGGGAGCTGGCTCAATGGAAGAGATGCAAAATGTTGATGTC	930
QY	830	TGCGCAGCCTTATCTAGAAAGTCCCATTTGCTTCTCCAAATAAATCCCTACAGAA	889
DB	931	ACGCCCTACCTTTGAGCGCGATTTGGCCAAITTCCTCAAAACAGCTCAACCTTTACCGGAT	990
QY	890	GGTCATTTGTTGGGTTGGTTGTTCTAAGCACTTCTGCTGCTACACGCTCTCAAAATCC	949
DB	991	AGTAATCAITTCCTGCTTATCATCTCTGCTCTCTCTTCTTCTCAATATCGTATCATCC	1050
QY	950	TGTCGCTAATGCAATACCACTGTCGCTTTTATCTGTTATATGAGATTTGGTTTGGCTT	1009
DB	1051	AGTCGCTAATGCTTATGAAATTTGGCTAGTATCTGTTATCTGTGAGGCTTGTTTGGCTT	1110

QY	1010	ATCCTGGATACTGATCAGTTCCCGAAGTGTCTTCCATCAACCGGAGACCTACTCTTGA	1069
DB	1111	GTCTTGGCTTCTAGATCAGTTCCCAAAATGATATCCATCAACCGTGAACATATCTCGA	1170
QY	1070	TAGACTGGCTTTAAGGTATGACCGAGAGGTGAACCGTCTCAGTTGGCTGTGTTGACAT	1129
DB	1171	CAGGCTTGCATTGAGGTATGATGAGAGGGAGAGCCATCAGAGTGGCTCCCATGATGT	1230
QY	1130	ATTGTCAGTACACTGACCCCTTGAAGGAGCCACTATCGTCTACTGCCCAACACTGTGCT	1189
DB	1231	CTTTGTGAGTACAGTGCATCCATTGAAGGAACTCTCACTGATCAACAGCAACACTGTTT	1290
QY	1190	ATCCATTCTTGTCTGTGATATCCCTGGCAAGGTCTCTGCTATGTATGTATGTATGACG	1249
DB	1291	GTCCATTCTTGTGTGATTAACCTGTTGACAAAGTGTCACTGTATGTTCTGTATGATGG	1350
QY	1250	AGCTTCAATGTGCTCTTTGACGCAATTTGGCTGAGACTTCAGAGTTTGTAGAAATGGGT	1309
DB	1351	CTCAGCTATGTGCTGCTTTTGTAGTCTCTCTCTGAAACTGCCGAATTTGCTAGAAAGTGGT	1410
QY	1310	ACCAATTTGTGAAGAGTATGACATTTGAACCCAGAGCTCCCGAGTTTCTACTTTGCCGAA	1369
DB	1411	TCCCTTTGTAAGAGCAATATTGAACCAAGAGCTCCAGAAATTTTACTTTGCTCAAAA	1470
QY	1370	AAATTGATTACTGAAAGACAAAGTCCAGCCTTCAATTTGTTAAAGACCGCGGGCCATCAA	1429
DB	1471	AATAGATTACTGAGGACAAATTTCAACTTCACTTTGTTAAAGAAAGACGAGCAATGAA	1530
QY	1430	GAGAGATATGAAGAAATTTAAATCAGGATAAATGCCCTAGTTTCTTAAGGCAATGAAAGT	1489
DB	1531	GAGAGATATGAAGAAATTTCAAAATTAAGATCAATGCCCTTGTGCCAAAGCACAGAAAGT	1590
QY	1490	CCCCGAGGAGGATGGAATGCAAGATGACACCATGCGCAGGAAACATATACACGGA	1549
DB	1591	GCCTGAAAGAGGGTGGACCATGGCTGATGGAATGCTTGGGCTTGGGAATTAACCTTAGGGA	1650
QY	1550	TCATCTCGGAATGATTCAGGTTTTTCCVTGGTTCACAGTGGTGGCTTGTATCTAGAGGTAA	1609
DB	1651	CCATCTCGCATGATTCAGGTTGTTCTTGGGGCACAGTGGTGGGCTTGCACACTGATGAAA	1710
QY	1610	TGAGCTCCCGCTTTAGTTTATGTTCTGTAAGAAAGCGTCTGGGTTTCAGACACCAAA	1669
DB	1711	TGAATTAACACGCTTCTGTCTATGTTCTGCGTGAAGAGAGACCGAGCTTTCAGCATCAA	1770
QY	1670	GAGGCTGTGTCATGAATCCCTTGTTCGTCTCAGCTGCTCTTACTACTATGACAAATA	1729
DB	1771	GAGGCTGGTGCATGAATGACCTGATTCGTGTATCTGCTGTGCTGCAAAATGTTGCTTA	1830
QY	1730	CATGTTGAATCTTTGATGATCACTACATCAACAAACAGCAAGGCTCTCCGAGAGGCTAT	1789
DB	1831	TCCTTCTCAATGTGGATTGTGACCATTAATCAATAGCAGCAAAAGCTCTTAGAGAGCAAT	1890
QY	1790	GTGCTTCTTAATGAGTCCAAACCTAGTCCGAAAGTCTGTTATGTGTCAGTTCCCAAAAG	1849
DB	1891	GTGCTTCATGATGATCCAGCTCTAGGAAAGGAAACCTTGTATGTACAAATTTCCACAAG	1950
QY	1850	GTTTGATGGATGATAGAGGAATGATGATGATGAAACAGGAAACACTCTCTTTTGTATAT	1909
DB	1951	ATTTGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2010
QY	1910	TAACTTGGAGGGCTTTGACCGCAATTCAGGACCAAGTTTATGTGGAACTGGTTGTGTTT	1969
DB	2011	CAACATGAAAGGCTAGATGCGATTCAGGGTCCAGTCTATGTGGGAACAGGATGCTGTTT	2070
QY	1970	CAACAGAACAGCTATCTATGTTTATGAGCC---CCCAATTAAGCGGAGAGCCAGGTTT	2026
DB	2071	CAATAGCAGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2130
QY	2027	CTTGGCATCACTATGTGGGGCAAGAGGAAAGCAAGTCAAGGAAAGAGGAGCTCAGA	2086
DB	2131	CATTGTTGTTAAGAGCTGCTGTGTAGAGGAGAGGAAAGAAACAGAGTTTATGATGATG	2190
QY	2087	TAAGAAAAAGTCGAAACAGCATGTGGACAGTTCTGTTCAGTATTCAATCTCGAAGACAT	2146

```
Db 2191 TCAGACCGCATATATAGAGAGACAGAACTCTTCAGCTCCCACTTTTAACTAGGAAGACAT 2250
QY 2147 AGAGAGGGGTGTGAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTCATGTCTCAAAAT 2206
Db 2251 CGAGGAGGGGTATTGAAGGT-----TATGAGGATGAAAGGTCAGTGCTTATGTGCCAGAG 2304
QY 2207 GAGCTTAGAGAGAGAGATTTGGCCAGTCAGCAGCAATTTGTGCTCCACCTCTGATGGAATA 2266
Db 2305 GAAATGGGAGAACGCTTTGGTCAGTCTCCAATCTTCATTGCAATCCACCTTTATGACTCA 2364
QY 2267 TGGTGTGTCTCTCAGTCTCCTCACTCCAGAACTCTTTTGAAGAAGCTATCCATGTCAAT 2326
Db 2365 AGTGGCATACACCTTCAACAAACCAGCTCTCTACTGAAGAGCTATCCAATGTTAT 2424
QY 2327 AGTTGTGCTATGAGGACAAGTCTGAATGGGAACTGAGATGGTGGATCTATGGATC 2386
Db 2425 CAGCTGTGGGTACGAGGACAAAACCTGAATGGGAAAAGAGATTGGCTGGATCTATGGTTC 2484
QY 2387 TGTCAAGAAGATATCTTACTGGATTCAAGATGACGCAAGAGGCTGGCTTCAGTCTA 2446
Db 2485 AGTTACAGAGGATATTTCTGACTGGGTTTAAATGCAATGCAAGAGGCTGGCAATCAATCTA 2544
QY 2447 TTGCATGCCCAAGCGCCAGCTTCAAGGATCTGCCCATCAATCTTTTCAGATCGTCT 2506
Db 2545 CTGCATGCCACACGACCTTGTTCAGGGTCTTCACCAATCAATCTTTCAGTGTCT 2604
QY 2507 GAACCAAGTGTGCGGTGGGCTCTCGGTTCTGTGTGAAATCTTTTTCAGCGGCAATGGCC 2566
Db 2605 TAATCAGGTGCTCCGTGGGCTCTTGGGTCAAGTGGAAATCTGCTTAGCAGACATTTGTC 2664
QY 2567 CTTATGGTATGCTACGGAGGGGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACAC 2626
Db 2665 TATATGGTATGGCTACAATGGGCGATGAAGCTTTTGGAGAGCTGGCTTACATTAACAC 2724
QY 2627 CACCAATTTACCACTAACTCTCTCCGCTCTTAGTCTATTTGATATTCGCTGCTATCTG 2686
Db 2725 CATTTGTTATCCAAATCAATCTGTTCCGCTTATCGCTATTTGTCGTTCTGCTATCTG 2784
QY 2687 TCTGCTCACTGGAAGTTCAATCAAGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTCAAT 2746
Db 2785 TCTTCTTACCAATAAAATTTATCATCTCTGAGATTAGTAATATGCTGGAATGTTCTTCAT 2844
QY 2747 TGGGCTCTTCCTTCAATTTTGGCACTGGTATCCTTGAGATCAGTGAAGTGGTGG 2806
Db 2845 TCTTCTTTTGGCTCCATTTTGCACATGATATTTGGAGCTCAGATGAGTGGTGG 2904
QY 2807 CATTCAGAGTGGTGGAGGAATGAACAGTTTCTGGGTCAATGGAGGTATCTTGCACATCT 2866
Db 2905 CATTGAAGATTGGTGAAGAAATGAGCAGTTTGGGTTATTTGGGCACCTCTGCCCATCT 2964
QY 2867 GTTTGCGGCTTTACGGGTCTTCTGAAGTGTCTTGGCGGTATGACACCAACTTCACTGT 2926
Db 2965 CTTGCGGTGTTCACGGGTCTGCTGAAGTGTGGCTGGATTGATACCAACTTCAAGT 3024
QY 2927 CACCTCAAGGCTAATGACGAAGAGCGACATTTGCTGAGCTCTACATGTTCAAGTGGAC 2986
Db 3025 TACCTCAAGGCTAATGATGAGATGGCAGCTTTGCTGAGCTATATGTTTCAAGTGGAC 3084
QY 2987 GAGCGTCTCTCATCCCTCCGACGACCAATTTGATCAATTAACATGGTGGTGGTCTGCTGG 3046
Db 3085 CAGTTTGTCTATCCCTCCGACCACTGTTCTTGTCTATTAACCTGGTCGGAATGGTGCAGG 3144
QY 3047 CACCTCTACGCAATCAACAGTGGTTTACCAATCATGGGCGCGCTTTTGGGAAGCTCTT 3106
Db 3145 AAATTCGTATGCCATTAACAGCGGCTACCAATCCTGGGGTCCGCTCTTTTGGAAAGCTGT 3204
QY 3107 CTTTGGCTTCTGGGTCAATTTGTTCACTTATACCAATTCCTCAAGGCTTATGGGAGGCA 3166
Db 3205 CTTCTCGATCTGGGTGATCCTCCATCTCTACCCCTTCTCAAGGGTCTCATGGGAGGCA 3264
QY 3167 AAACCGCACACCGAATTTGTCATCGTCTGGGCTGTCCTCGCTTCTATCTTCTCTT 3226
```

```
Db 3265 GAACCGCAGCCCAACAACTGTCATCGTTGGTCCATCCTTGGCTATCTTCTCCTT 3324
QY 3227 GCTGTGGGTTGCTGTGATCCATTCACCTACCC 3258
Db 3325 GCTGTGGGTGAAGATCGATCCTTTCATCTCCC 3356

RESULT 12
AAZ99521
ID AAZ99521 standard; DNA; 3725 BP.
XX
AC AAZ99521;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
FH Key Location/Qualifiers
FT CDS 179..3397
FT /tag= a
FT /product= "cellulose synthase"
FT /note= "no termination codon given"
XX
PN W0200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
PR 17-AUG-1998; 98US-0096822P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
DR P-PSDB; AAY84117.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
XX
Claim 1; Page 160-165; 119pp; English.
XX
The present sequence encodes a maize cellulose synthase polypeptide. The
cellulose synthase can be used for the improvement of stalk quality for
improved stand or silage. It also provides an increased concentration of
cellulose in the pericarp, hardening the kernel and improving its
handling ability. The sequences are used to produce transgenic plants and
seeds expressing the cellulose synthase. The polynucleotide is used for
modulating, preferably increasing, the level of the synthase in a plant
cell. The plants are preferably monocots. The polynucleotide is also used
as a probe or primer in the detection quantitation or isolation of gene
transcripts. The probes are useful in detecting deficiencies in the level
of mRNA in screenings for desired transgenic plant, for detecting or
mutations in the gene, for monitoring upregulation of expression or
changes in enzyme activity in screening assays of compounds, for
detection of any number of allelic variants of the gene, or for use as
molecular markers in plant breeding programs. The isolated nucleic acids
of the present invention can also be used for recombinant expression of
their encoded polypeptides or for use as immunogens in the preparation
and/or screening of antibodies. The proteins can be employed in assays
for enzyme agonists or antagonists of enzyme function or for use of
immunogens or antigens to obtain antibodies specifically immunoreactive
with a protein
XX
Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 U; 0 Other;
```

Query Match	38.8%;	Score 1408.4;	DB 3;	Length 3725;	
Best Local Similarity	66.8%;	Pred. No. 0;			
Matches 2147;	Conservative	0;	Mismatches 966;	Indels 99;	Gaps 6;
QY	53	CGCGCCATGGACGGCGACGCGAGCCGCTGAAGTCCGGGAGGCACGCGGCGCGGGACGT	112		
Db	238	CGCCACGACGGGACGCGCTGTCCCGCTAAGCCACGAGAGTGGCAATGGGACAGT	297		
QY	113	GTGCCAGATCTCGCGCGACGGCCTGGGCACCAAGTGTGGACCGGACGCTTTCAACCGCTG	172		
Db	298	CTGCCAGATTTGTGGCGACACACTGTTGGCGTTTCAGCCACTGGTAGTCTTTGTGCGTG	357		
QY	173	CGAGCTCTCGCTTCCCGCTCTCGCCCTCTGCTACAGACGCGACGCGCGAGGAGGCAC	232		
Db	358	CAATGAGTGTGCTTCCCTGTCTGCGCCCTTGTCTATGAGTACGAGCGCAAGGAGGAA	417		
QY	233	CCAGGCTCGCTCCAGTGCACAGCAAGTACAAGCGCCACAGAGGGAGCCGACGATCCG	292		
Db	418	CCAAATGCTCCCTCAGTGCAGACTAGATACAAGAGACAGAAAGTACGCCCTCGAGTTCA	477		
QY	293	CGGGAGGAGGCGACGACACTGATGCCGATGATGGTAGTCACTTCAACTACCTTGCATC	352		
Db	478	TGGTGAATGATGAGGAGGAAGATGTTGATGACTGGACAATGAATTCAACTAATAGCAAGG	537		
QY	353	TGGCACTGAGGACCAAGACGAAAGATTGCTGACAGGATGGCGAGCTGGCGCATGAACAC	412		
Db	538	CAATGGNAGGGCCAGAGTGGCAGCTT-----	574		
QY	413	CGGGGCGATGGCAATGTTGGCCACCCCAAGTATGACAGTGGGAGATCGGCCCTCTCCAA	472		
Db	575	GAGCGTGATCTGTTCTATCTCTCGCCATGACCCACACCAATCGGATTTCCACGCTTACA	634		
QY	473	GTATGACAGTGGAGATCCCTAGGGGATACGTCCCTTCAGTCAACACGCCAGATGTC	532		
Db	635	AGTGGACA-----	651		
QY	533	AGGAGAAATCCCTGGAGCTTCGCTGATCATCACATGATGCCCTACGGGAAACATCAG	592		
Db	652	TGGAGAGATCCCTGATGATCCCTGACCGTCAATCTATCCGAGTCAACATCGAGCTA	711		
QY	593	CAGACGTGCTCGTTTCCCTATGTGAATCATCCCAATCCGTCAAGGAGTT---CTC	649		
Db	712	TGTTGATCAAGGTTTCCAGTTCCTGTGAGAGATTGGACCCCTCGAAGGACTTGAATTC	771		
QY	650	CGSCAGTATGGGAATGTTGCTCGTGAAGAGAGAGATTGATGGCTGGAAATGAAGCAGGA	709		
Db	772	CTATGGGCTTAATAGTGTGACCTGGAAGGAAAGAGTTGAGAGCTGGAGGGTTAAACAGGA	831		
QY	710	CAAGGTGGATCCCATGACATAATGGGCAAGCATGCTCGCTCTGAAGTGGGCGACG	769		
Db	832	CAAAAAATATGTTGCAAGTACTAA-----	870		
QY	770	TACTGACATCGATGCAATCTACTGAATCAAAATGAAGACGCTTTACTGATGATGAAC	829		
Db	871	TAGAGGAGACATCGAGGAGCTGGCTCAATGAGAGATATGCNAATGGTTGATGATGC	930		
QY	830	TCGCCAGCTCTATCTAGAAAAAGTCCCATGCTTCTCCAAATAAATCCCTACAGAAAT	889		
Db	931	ACGCCCTACCTTTGAGCGCATTTGTGCCAATTTCTCAAAACACAGCTCAACCTTTTACCGGAT	990		
QY	890	GGTCAATGTTCTGCGGTTGTTGTTCTAAGCATCTTCTGACACTTCTCGACTACCGTCTCAAAATCC	949		
Db	991	AGTAAATCTCTCGGCTTATCATCTGTGCTTCTTCTTCCAAATCGTATCAGTCAATCC	1050		
QY	950	TGTCGTAAATGCAATCCCACTGTGCTTTTATCTGTTATATGTGAGATTTGGTTTGTCTTT	1009		
Db	1051	AGTGCCTAATGCTTATGGAATTTGGCTAGTATCTGTTATCTGTGAGGCTCTGTTTGCCTT	1110		
QY	1010	ATCCTCGATACCTGATCTCCGAAAGTGGTTTCCATCAACCGGGAGACCTACCTTGA	1069		
Db	1111	GTCTCGCTCTTAGATCTAGTCTCCCAAAATGGTATCCAAATCAACCGTGAGACATATCTCGA	1170		
QY	1070	TAGACTGGCTTTAAGGATATGACCGGAGAGGTGAACCGCTCTCAGTTGGCTGCTGTTGACAT	1129		

Db 2251 CGAGGAGGATATTGAAGT-----TATGAGGATGAAGGTCAGTCTTATGTGCCAGAG 2304
QY 2207 GAGCTTAGAGAGAGATTGGCCAGTCAGCAGCAGATTGTTGGCTCCTCCTCATCTCTGATGAATA 2266
Db 2305 GAAATTTGGAGAACGCTTTGGTCAGTCTCCAATCTTCATTGTCATCCACCTTTATGACTCA 2364
QY 2267 TGGTGGTGTCTCAGTCTCCACTCCCAATCTCTTTTGAAGAGAGCTATCCATGTCTAT 2326
Db 2365 AGTGGCAATACCACTTCAACAAACCCAGCTCTCTACTGAAGGAGCTATCAATGTAT 2424
QY 2327 AAGTTGTGCTATGAGCAGCAAGTCTGAATGGGAACTGAGATTGGTTGGATCTATGGATC 2386
Db 2425 CAGCTGTGGTACGAGGACAAACTGAATGGGAAAAGAGATTGGCTGATCTATGGTTTC 2484
QY 2387 TGTACAGAGATATTTCTACTGGATTCAAGATGACGGAAGAGCTGGCTTCTAGTCTA 2446
Db 2485 AGTTACAGAGGATATTTCTACTGGGTTTAAATGTCATGAAGAGGCTGGCAATCAATCTA 2544
QY 2447 TTGCATGCCCAAGCGCCAGCTTCAAGGGATCTGCCCCCATCAATCTTTTCAGATCGTCT 2506
Db 2545 CTGCATGCCCAAGCAGCTTGTCTTCAAGGGTCTTCACCAATCAATCTTTCTGATCGTCT 2604
QY 2507 GAACCAAGTGTGCGGTGGGCTCTCGGTTCTGTGAAATCTTTTCAGCGGCATTGCC 2566
Db 2605 TAATCAGGTGCTCGTGGGCTCTGGGTCAGTGGAAATCTGCTTAGCAGACATTTGTCC 2664
QY 2567 CTTATGGTATGCTACGAGGGGCGCTCAAGTTCTCGGAGAGATTCGCTTATCAATCAAC 2626
Db 2665 TATATGGTATGCTACAAATGGGCGATTGAAGCTTTTGGAGAGGCTGGCTTACATTAAC 2724
QY 2627 CACCAATTTACCACTCACTCTCCGCTCTAGTCTATTTATATTGCTTCCCTGCTATCTG 2686
Db 2725 CATTGTTTATCAATCAATCAATCTGCTTCCGCTATTCGCTTATGCTTCTGCTATCTG 2784
QY 2687 TCTGCTCACTGGAAGTTCATCATGCCAGAGATTAGCAACTTGGCCAGATCTCTGCTTAT 2746
Db 2785 TCTCTTACCAATAAATTTATCTCTGAGATTAGTAATATGCTGGAATGTTCTTCAT 2844
QY 2747 TCGGCTCTTCTTCAATTTTGGCCACTGCTGATCTTGTGAGATGAGTGGAGTGGTTGG 2806
Db 2845 TCTTCTTTTGGCTCCATTTTGGCACTGCTATTTGGAGCTCAGATGGAGTGGTGGTGG 2904
QY 2807 CATTCAGAGTGGTGGAGGAATGAACAGTTCTGGGTCATTGGAGGATCTCTGACATCT 2866
Db 2905 CATTGAAGATTGGTGAGAAATGAGCAGTTTGGGTTATGGTGACCTCTGCCCATCT 2964
QY 2867 GTTTCGGCTCTTTCAGGGTCTTCTGAAGTGTCTTGGCGTATCGACACCAACTTCACTGT 2926
Db 2965 CTTGGCGGTGTTCCAGGGTCTGCTGAAAGTGTGGCTGGGATTGATACCAACTTCACGT 3024
QY 2927 CAGCTCAAAGGCTAATGACGAGAGGAGCATTGCTGAGCTCTACATGTTCAAGTGGAC 2986
Db 3025 TACCTCAAAGGCTATGATGAGGATGGGACTTGTGAGCTATATGTTGTTCAAGTGGAC 3084
QY 2987 GAGCTTCTCATCCCTCCGAGCACCATTGATCAATGATGGTGGTCTGCTGCTGG 3046
Db 3085 CAGTTTGTCTATCCCTCCGACCACTGTTCTGTCAITTAACCTGGTCGGAATGGTGCAG 3144
QY 3047 CAGCTCTACGCCATCAACAGTGGTTTACCAATCATGGGGCGCGCTCTTTGGGAAGCTCTT 3106
Db 3145 AATTTGATGATCCATTAAACAGGGCTTACCAATCTCTGGGGTCCGCTCTTTGGAAAGCTGTT 3204
QY 3107 CTTTGGCTTCTGGGTGATGTTGACCTTATACCATTCCTCAAGGCTCTTATGGGAGGCA 3166
Db 3205 CTTCTCGATCTGGGTGATCTCCATCTCTACCCCTTCTCAAGGGTCTCATGGGAGGCA 3264
QY 3167 ABAACCCACACGACGATGTCATCTCTGGGCTGCTCCCTCTCTATCTTCTTCCCTT 3226
Db 3265 GAACCCGACGCAACAAATGCTATCGTTTGGTCCATCCCTCTTGGCGTCTATCTTCTCCTT 3324
QY 3227 GCTGTGGGTTCGTTGTTGATCCATTCATACCC 3258
Db 3325 GCTGTGGGTGAAGATCGATCTTTCATCTCCC 3356

RESULT 13
AAZ99506
ID AAZ99506 standard; DNA; 3725 BP.
XX
AC AAZ99506;
XX 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
FH Key Location/Qualifiers
CDS 179..3397
FT /*tag= a
FT /product= "cellulose synthase"
FT /note= "no termination codon given"
XX
XX WO200009706-A2.
XX
XX 24-FEB-2000.
XX
XX 16-AUG-1999; 99WO-US018760.
XX
XX 17-AUG-1999; 98US-0096822P.
XX
XX (PION-) PIONEER HT-BRED INT INC.
XX
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
XX WPI; 2000-224343/19.
XX P-PSDB; AAY84112.
XX
XX New genes which encode maize cellulose synthase polypeptides in plants
XX useful for modulating the expression of cellulose synthase in plants and
XX to produce transgenic plants expressing the novel protein.
XX
XX Claim 1; Page 121-126; 119pp; English.
XX
XX The present sequence encodes a maize cellulose synthase polypeptide. The
XX cellulose synthase can be used for the improvement of stalk quality for
XX improved stand or silage. It also provides an increased concentration of
XX cellulose in the pericarp, hardening the kernel and improving its
XX handling ability. The sequences are used to produce transgenic plants and
XX seeds expressing the cellulose synthase. The polynucleotide is used for
XX modulating, preferably increasing, the level of the synthase in a plant
XX cell. The plants are preferably monocots. The polynucleotide is also used
XX as a probe or primer in the detection quantitation or isolation of gene
XX transcripts. The probes are useful in detecting deficiencies in the level
XX of mRNA in screenings for desired transgenic plant, for detecting
XX mutations in the gene, for monitoring upregulation of expression or
XX changes in enzyme activity in screening assays of compounds, for
XX detection of any number of allelic variants of the gene, or for use as
XX molecular markers in plant breeding programs. The isolated nucleic acids
XX of the present invention can also be used for recombinant expression of
XX their encoded polypeptides or for use as immunogens in the preparation
XX and/or screening of antibodies. The proteins can be employed in assays
XX for enzyme agonists or antagonists of enzyme function or for use of
XX immunogens or antigens to obtain antibodies specifically immunoreactive
XX with a protein
XX
XX Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 U; 0 Other;

Query Match 38.8%; Score 1408.4; DB 3; Length 3725;
Best Local Similarity 66.8%; Pred. No. 0;
Matches 2147; Conservative 0; Mismatches 966; Indels 99; Gaps 6;

QY 53 CCGCGCCATGGACGGGACGCGGACGCCCTTGAAGTCCGGGAGGCACGGGGCGGGGACGT 112

Db 238 CCGCCACGACGGCGCGCTGTCGGCTAAGCCACGAGAGTGGCAATGGCGAGT 297
Qy 113 GTGCCAGATCTGCGCGACGGCCCTGGGACCACTGTTGGAACGCGACGCTTTCAACGCGTG 172
Db 298 CTGCCAGATTTGTGGCGACACTGTGGCGTTTCAGCCACTGGTGATGCTTTGTGTGCGTG 357
Qy 173 CGAGCTCTGCGGCTTCCGGTCTCGCGCCCTGTACGACACGACGCGCAGGAGGCGAC 232
Db 358 CAATGAGTGTGCTTCCCTGCTCGCGCCCTGTATGAGTACGAGCGCAAGGAAGGAA 417
Qy 233 CCAGGCTGCTCCAGTGCACAGACCAAGTACAAGCGCCACAGAGGAGCCCGACGATCCG 292
Db 418 CCAATGCTGCCCTCAGTGCACAGCTAGTACAAGAGACAGAAAGTAGCCCTCGAGTTCA 477
Qy 293 CGGGAGGAGGACGACACACTGATGCGGATGATGATGATGATGATGATGATGATGATG 352
Db 478 TGGTGATGATGAGGAGGAAGATGTTGATGACCTGGCAATGAATCAACTATAAGCAAGG 537
Qy 353 TGGCACTGAGGACGAGGAGGAAGATGCTGACAGGATGCGAGCTGGGCGCATGAACAC 412
Db 538 CAATGGGAAGGCCGACAGTGGCAGCTT-----CAAGGAGAT 574
Qy 413 CGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCTCTCCAA 472
Db 575 GACGCTGATCTGCTTCACTGCTCGCATGACCCACACCATCGAATTCACAGGCTTTACA 634
Qy 473 GTATGACAGTGGAGAGATCCCTAGGGGATAGCTCCCTTCAGTACCAACAGCGCAGATGTC 532
Db 635 AGTGGACA-----ACAGATATC 651
Qy 533 AGGAGAAATCCCTGGAGCTTCGCTGATCATCATGATGTCCTACGGGGAACATCAG 592
Db 652 TGGAGAGATCCCTGATGATCCCTGACCGCTCATCTATCGGAGTCAACATCGAGCTA 711
Qy 593 CAGACGTCTCGGTTCCCTATGTAATCATCCAAATCCGTCACAGGGAGTT---CTC 649
Db 712 TGTGTATCCAAAGCGTTCAGGTTCTGTGAGGATTTGACCCCTCGAAGGACTTTGAATTC 771
Qy 650 CGGCGATTTGGGAATGCTGCTGGAAAGAGAGAGTGTGAGCTGGGAAATGAAGCAGGA 709
Db 772 CTATGGGCTTAATAGTGTGACTGGAAGGAAGAGTTGAGAGCTGGAGGGTTAAACAGGA 831
Qy 710 CAAGGGTGCCATTCCTCATGACTAATGGAACAAGCATTTGCTCCCTCTGAAGTGGGCGAC 769
Db 832 CAAAAATATGTTGCAAGTGACTAA-----TAAATATCCAGAGC 870
Qy 770 TACTGACATCGATGCTACTACTGAATACAACATGGAAGACGCTTTACTGATGATGAAC 829
Db 871 TAGAGGAGACATGGAGGGGACTGGCTCAATGGAGAAGATATGCAATGGTTGATGATGC 930
Qy 830 TCGCCAGCTCTATCTAGAAAGTCCCATGCTTCTCTCCAAATAAATCCCTACAGAT 889
Db 931 ACGCTACCTTTGAGCGCGCATTTGCGCAATTTCTCAACACGCTCAACCTTTACCGGAT 990
Qy 890 GGTCAATGTTCTCGGTTGTGTCTTAAGCATCTTCTGCACTACCGTCTCAAAATCC 949
Db 991 AGTAATCATCTCGCTTATCATCTCTGCTGCTTCTTCTTCCAATATGATCAGTCAATCC 1050
Qy 950 TGTGCGTAATGCTATACCACTGTGGCTTTTATCTGTTATATGAGAGATTTGGTTTGTCTT 1009
Db 1051 AGTGCCTAATGCTTATGGAATGTGGCTAGTATCTGTTATCTGTGAGGCTTGGTTTGCCTT 1110
Qy 1010 ATCTGATGACTGATCAGTCCGAGTGGTTCCAAATCAACGGGAGACCTTACCTTGA 1069
Db 1111 GTCTGGCTCTTAGATCAGTTCCAAATGGTATCCAAATCAACCGTGAGACATATCTCGA 1170
Qy 1070 TAGACTGGCTTTAAGGTATGACCGAGAGGAGTGAACCGTCTCAGTTGGCTGCTGTTGACAT 1129
Db 1171 CAGGCTTGCATTTAGGATGATGAGAGGGGAGAGCCATCAAGCTGCTCCCATTTGATGT 1230
Qy 1130 ATTGTGATGATGACCGCTTTGAAGGAGGACCTATCGTCACTGCCAACACTGTGCT 1189

Db 1231 CTTTGTGATCAGTACAGTGGATCCATTTGAAGAACCTTCCACTGATCACAGCCACACTGTTTT 1290
Qy 1190 ATCCATTTCTTGTGCTGTTGATATCCGTTGGACAAGGTCTCTTGTGATGATCTGATGACGG 1249
Db 1291 GTCCATTTCTTGTGCTGTTGATACCTGTTGACAAAGTGTCTGCTATGTTCTGATGATGG 1350
Qy 1250 AGCTTCAATGCTGACTTTTTCAGCATGCTGGCTGAGACTTCAGAGTTTCTAGGAATGGGT 1309
Db 1351 CTCAGCTATGCTGACTTTTGTGATCTCTCTGAAACTGCCGAATTTGCTAGAAAGTGGGT 1410
Qy 1310 ACCATTTGTGAAGAAGTATGACATTTGAACCCAGAGCTCCCGAGTTTTACTTTTCCAGAA 1369
Db 1411 TCCCTTTTGAAGAAGCACAAATTTGAACCAAGAGCTCCAGAAATTTTACTTTTGTCTCAAA 1470
Qy 1370 AATTGATCTCTGAAAGACAAAGTCCAGCCTTCAATTTGTTAAAGACCGCGGGCCATGAA 1429
Db 1471 AATAGATTACTGAAAGGACAAATTTCAACCTTCATTTTGTAAAGAAAGACGAGCAATGAA 1530
Qy 1430 GAGAGAAATATGAAGAAATTTAAATTCAGATAAATGCCCTAGTTTCTTAAGGCATTTGAAAGT 1489
Db 1531 GAGAGAGTATGAAGAAATTCAAATTAAGATCAATGCCCTTGTTCGCAAGACACAGAAAGT 1590
Qy 1490 CCGCAGGAAGGATGGATCATGCAAGATGGCACACCATGGCCAGGAACAATACAGGGA 1549
Db 1591 GCCTGAAGAGGGGTGGACCATGGCTGATGAACTGTCTGGCTGGGAAATAACCCCTAGGGA 1650
Qy 1550 TCATCTTGGAAATGATTCAGGTTTTCTTGTGTCACAGTGGTGGCTTGTATCTAGAGGTAA 1609
Db 1651 CCATCTCTGGCATGATTCAGGTTCTTTGGGGGACAGTGGTGGGCTTGACACTGATGGAAA 1710
Qy 1610 TGAGCTCCCCCGTTTGTGTTTATGTTCTCTGTAAGAACGCTCTGGGTTTCAGACACCAAA 1669
Db 1711 TGAATTAACAGCTCTTGTCTATGCTCTCTGTAAGAGAGACAGGCTTTTACGATCAAA 1770
Qy 1670 GAAGGCTGGTCCCATGAATCCCTTGTGTTCTGCTCTCAGCTGTCTTACTAATGGAACAATA 1729
Db 1771 GAAGGCTGGTGCATGAATGAATGCACTGATTCGTGATCTGCTGTGTAACAATGGTGCTTA 1830
Qy 1730 CATGTTGAATCTTGTGTTGATCATTACTATCATATCAACAGAGGCTCTCGAGAGCTAT 1789
Db 1831 TCTTCTCAATGTGGATTGTACCAATTACTTCAATAGCAGCAAGCTCTTAGAGAAGCAAT 1850
Qy 1790 GTGCTTCTCAATGGATCCAAACCTTAGGTCGCAAGTCTGTTATGTGCAAGTTCCTCCACAAAG 1849
Db 1891 GTGCTTCTCATGATGATCCAGCTCTAGGAAGAAACTTGTATGTAACAATTTCCACAAG 1950
Qy 1850 GTTTGATGGGATGATAGGAATGATCGATATGAAACAGGAACACTGCTGTTTTTTTGTATAT 1909
Db 1951 ATTTGATGGCATTTGACTTGCACGATCGATATGCTTAATAGGAACATAGTCTTCTTTGATAT 2010
Qy 1910 TAACTTGAGGGGCTTGCACGCTTCAAGGACAGTATTATGTGGGAACCTGGTTGTGTTTT 1969
Db 2011 CAACATGAAGGTTCTAGATGGCATTCAGGTTCCAGTCTATGTGGGAACAGATGCTGTTTT 2070
Qy 1970 CAACAGAAACAGCTCTATGTTTATGAGCC---CCCAATTAAGCGCAAGACCGAGTTT 2026
Db 2071 CAATAGCAGGCTTTGTATGATATGATCTGTTTGTGCTGAGCTGATCTGGAACCTAA 2130
Qy 2027 CTTGGCATCATGATGTTGGGGCAAGAGAGGCAAGCAAGCAAGTCAAGAAAAGGAGCTCAGA 2086
Db 2131 CATTTGTTTAAAGAGCTGCTGTGTAGGAAGAGAGAAAGAAAGAGTTATATGGATAG 2190
Qy 2087 TAAGAAAAGTCCGAACAAGCATGTGGACAGTCTGTGTTCCAGTATTCATCTCGAAGACAT 2146
Db 2191 TCAAGCCGTATTATGAAGAGACAGAAATCTTCACTGCCATCTTTAATGGAAGACAT 2250
Qy 2147 AGAGGAGGGTGTTCAGAGGCTGCTGGGTTTGTATGATGAGAAATCAGTTCTCATGCTCAAA 2206
Db 2251 CGAGGAGGGTATTGAAGGT-----TATGAGGATGAAGGTCAAGTCTTATGTTCCAGAG 2304
Qy 2207 GAGCTTAGAAGAGAGATTGGCCAGTCAGCAGCATTTGTTGGCTCCACTCTGATGGAATA 2266
Db 2305 GAAATTTGGAGAAAACGCTTTGGTCACTCCAAATCTTCAATTCATTCACCTTTTATGACTCA 2364

```
QY 2267 TGGTGGTGTCTCTCAGTCTCTCACTCCAGAACTCTTTTGAAGAGAGCTATCCATGCTAT 2326
Db |||||
QY 2365 AGGTGGCATACACCTTTCAACAACCCAGCTTCTTACTGAAGAGAGCTATCCATGTTAT 2424
Db |||||
QY 2327 AAGTTGTGGCTATGAGGACCAAGTCTGAATGGGAACTGAGATTGGTGGATCTATGGATC 2386
Db |||||
QY 2425 CAGCTGTGGTAGGAGACAAACTGAATGGGAAAGAGATGGCTGGATCTATGGTTC 2484
Db |||||
QY 2387 TGTACAGAGATATTTCTTACTGGATTCAAGATGCGACGAGAGGCTGGCTTCAGTCTTA 2446
Db |||||
QY 2485 AGTTACAGAGGATATTTCTGACTGGTGTAAATGSCATGCAAGAGGCTGGCAATCAATCTA 2544
Db |||||
QY 2447 TTGCATGCCCAAGCGCCAGCTTTCAAGGATCTGCCCATCAATCTTTTCAGATCGTCT 2506
Db |||||
QY 2545 CTGCATGCCACACGACCTTGTTCAGGGTCTGACCAATCAATCTTTCTGATCGTCT 2604
Db |||||
QY 2507 GAACCAAGTCTGCGTGGGCTCTCGGTTCTGTGAAATTTCTTTTCAGCGCGCAITGGCC 2566
Db |||||
QY 2605 TAATCAGGTGCTCGGTTGGGCTCTTGGGTCAGTGGAATTTCTGTTAGCAGACATTTGTC 2664
Db |||||
QY 2567 CTTATCGTATGGCTACGAGGGGCGCTCAAGTTCCTGAGAGATTCGCTTACATCAACAC 2626
Db |||||
QY 2665 TATATGGTATGGCTACAATGGGCGATTGAAGCTTTTGGAGGCTGGCTTACATTAACAC 2724
Db |||||
QY 2627 CACCATTACCCACTAACCTCTCTCCGCTTCTAGTCTATTTGATATTGCCTGCTATCTG 2686
Db |||||
QY 2725 CATGTTTATCAATCAATCAATCTGTCGCTTATCGGCTATTTGTCCTTCTGCTATCTG 2784
Db |||||
QY 2687 TCTGCTCACTGGAAGTTTCATATGCGCAGAGATTAGCACTTGGCCAGTATCTGTTTCAT 2746
Db |||||
QY 2785 TCTTCTTACCAATAAATTTATCATCTCTGAGATTAGTAATATGCTGGAATGTTCTTCAT 2844
Db |||||
QY 2747 TGGGCTCTTCTTCAATTTTGGCACTGATATCTTGGATGAGTGGAGTGGTGG 2806
Db |||||
QY 2845 TCTTCTTTTGGCTCCTATTTGCAACTGATATATTGGAGCTCAGATGGAGTGGTGG 2904
Db |||||
QY 2807 CATTCAGAGTGTGGAGGAATGAACAGTTCCTGGGTCATTGGAGGATCTCTGCACATCT 2866
Db |||||
QY 2905 CATTTGAAGATTTGGGAGAAATGAGCAGTTTGGGTTATTTGGGCACTCTTGGCCATCT 2964
Db |||||
QY 2867 GTTTGGGCTCTTTTCAAGGCTCTTCTGAGGCTGCTGCGGCTATCGACCAACTTCACTGT 2926
Db |||||
QY 2965 CTTTGGGCTTCTTCCAGGGTCTGCTGAAAGTGTGGCTGGATGTATACCACTTCAAGT 3024
Db |||||
QY 2927 CACCTCAAGGCTATGACAGAGAGGCGACTTTGCTGAGCTTCAATGTTCAAGTGGAC 2986
Db |||||
QY 3025 TACCTCAAGGCTATGATGAGGATGGGCACTTTGCTGAGCTATATGTTTCAAGTGGAC 3084
Db |||||
QY 2987 GACGCTTCTCATCCCTCCGACGACCAATTTTGATCATTAACATGTTGGTGTGCTGCTG 3046
Db |||||
QY 3085 CAGTTTGTCTATCCCTCCGACCACTGTTCTTGTCTATTAACCTGTCGGAATGTTGCAGG 3144
Db |||||
QY 3047 CACCTCTTACGCCATCAACAGTGGTTTACCAATCATGGGGCGGCTCTTTTGGGAAGCTCTT 3106
Db |||||
QY 3145 AATTTGATGATGCAATTAACAGGGCTACCAATCTGGGGTCCGCTCTTTTGGAAAGCTGTT 3204
Db |||||
QY 3107 CTTTGGCTTCTGGGTAATTTGATCACTTATACCAATTCCTCAAGGCTTATGGGAGGCA 3166
Db |||||
QY 3205 CTTTCTGATCTGGGTGATCTCTCATCTCTACCCCTTCTCAAGGGTCTCATGGGAGGCA 3264
Db |||||
QY 3167 AAACGCGCACGACGATGTCATCGTCTGGGCTGCTCTTCTGCTCTTCTATCTTCTCTT 3226
Db |||||
QY 3265 GAACCGACGCGCAACAAATGTCATCGTTTGGTCCATCTCTTCTGCTCTATCTTCTCTT 3324
Db |||||
QY 3227 GCTGTGGGTTGCTGTTGATCCATTTCACTACCC 3258
Db |||||
QY 3325 GCTGTGGGTCAGATGATCCTTTTCATCTCCC 3356
Db |||||
```

RESULT 14

AAA67114

ID AAA67114 standard; DNA; 3851 BP.

```
XX AAA67114;
AC
XX 31-OCT-2000 (first entry)
DT
XX
DE Pinus radiata cellulose synthase nucleotide sequence SEQ ID NO:69.
XX
XX Eucalyptus grandis; pinus radiata; Monterey pine; modification;
KW plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
KW transgenic plant; ds.
XX
XX Pinus radiata.
OS
XX WO200022092-A2.
PN
XX 20-APR-2000.
PD
XX 08-OCT-1999; 99WO-NZ000169.
PF
XX 13-OCT-1998; 98US-00170862.
PR
XX 11-AUG-1999; 99US-0148426P.
PR
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
PA
XX Bloksberg LN;
PI
XX WPI; 2000-339328/29.
XX P-PSDB; AAB16307.
XX
XX New genes encoding proteins involved in a plant polysaccharide
PT biosynthetic pathway, useful for modulating or altering the
PT polysaccharide content, composition or structure of the plant.
XX
XX Claim 1; Page 71-72; 301pp; English.
XX
XX The present invention describes isolated polynucleotides (PN) comprising
CC a sequence selected from one of 835 nucleotide sequences given in
CC AAB67073 to AAB67907, their (reverse) complements, sequences producing an
CC Expectation (E) value of 0.01 or less compared to the 835 sequences,
CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
CC 835 sequences or sequences that are degenerately equivalent or allelic to
CC the 835 sequences. The polynucleotides are used to modify the activity of
CC a polypeptide involved in a polysaccharide biosynthetic pathway in the
CC plant. They are especially used to modulate or alter the polysaccharide
CC content, composition or structure of the plant. AAB16268 to AAB16340 are
CC proteins encoded by some of the polynucleotide sequence given in the
CC present invention
XX
SQ Sequence 3851 BP; 1040 A; 796 C; 944 G; 1071 T; 0 U; 0 Other;
Query Match 38.1%; Score 1382; DB 3; Length 3851;
Best Local Similarity 66.4%; Pred. No. 0;
Matches 2142; Conservative 0; Mismatches 1040; Indels 46; Gaps 9;
QY 75 GACGCCCTCACTCGGAGGACCGGGCCGGGAGCTGTCAGATCTGCCCGACGCG 134
Db |||||
QY 222 GAGGACCGCAAGCCTTTTGAACACGTTGAGTGCCAGCTCTGCCAGATTTGTGCGGAGGAC 281
Db |||||
QY 135 CTGGGCGACCACTGTGGACGGCGACGCTTTCACCGCTGCGAGCTCTGCCGCTTCCCGGTC 194
Db |||||
QY 282 GTGCGGCTTAAACACAGACGCGAGCTGTTCGTGCTGTAAAGTGTGCGGTTCTCTGTC 341
Db |||||
QY 195 TGCGCCCTCTGTCTAGAGCAACGAGGAGGAGGCCAGGCCCTGCTCCAGTGCAG 254
Db |||||
QY 342 TGTGCGGCTGTCTATGATACGAGACGAGAAGGAAATCAGTCGTCGCCGAGTCAAT 401
Db |||||
QY 255 ACCAAGTACAAGCGCCACAGAGGAGGCCAGCGATCCGCGGAGGAGGAGGACGACACT 314
Db |||||
QY 402 ACTCGTTACAAGCGTCAAAAGGGAGTCCA-----CGGGTGGAAAGGTGACCATGATAA 455
Db |||||
QY 315 GATGCGGATGATGGTAGTGAATTCACCTACCTGCTGCGACTTGGCACTGAGGACGAGAGCAG 374
Db |||||
```

Db 456 GAAGACGCTTGATGACATAGAACATGAATTTAATGTGGAGACTCAGCAAAAGAAAACAGGCG 515
Qy 375 AAGATTGCTGACAGGATCGGAGCTGGCGCATGAACACCGGGGCGAGTGGCAATGTTGGC 434
Db 516 CAGATCACCGAGGGGATGCTCCACGGACGATGAGCTATGGCCGAGGTCCCGACGACGAA 575
Qy 435 CACCCCAAGTATG-ACAGTGGCGAGATCGGCTCTCCA-AGTATGACAGTGGAGAGATCC 492
Db 576 AATTGGCAGATTGCTCATATTCAGAGTTCCTCCGACAGATTCCTGTACTTGGCAACGGC 635
Qy 493 CTAGGGATACGTCCTTCAGTCCAAACAGCCAGATGTCCAGAGAAATCCCTCGAGCTT 552
Db 636 CACTCGGTTGTAGTGGGAGATTCCAAACGTCTACTACGCAGCAACAACCAATTG----- 689
Qy 553 CGCCTGATCATCATGATGTCCTCTAGGGGACATCAGCAGACGTGCTCCGTTCCCT 612
Db 690 ---CTTGCCAACCTGCAATGCTGAAGCGTGTGCATCCAAGTCCGAGCCGGGAGTGA 746
Qy 613 ATGTGAATCATTCACCAAAATCCGTCAAGGGAGTTCTCCGGCAGTATTTGGGAATGTTGCC 672
Db 747 AGGATCATGATGATCCAAACAGGATATTTGTTCTTATG--GCTTTGGGAACGTGCTT 804
Qy 673 GGAAGAGAGAGTGTGATGCTGGAATGAAGCAGAGCAAGGGTGCATTCCTCATGACTA 732
Db 805 GGAAGGAGCGAGGCGATGGTTATAAATCGAAGGAAAACAAATCAGGCCAGTTGGATATGA 864
Qy 733 ATGGGACAGATTGCTCCCTCTGAAGTGGGCGAGCTACTGACATCGATGCATCTACTG 792
Db 865 CGGAAGGAGATATCAATATATGAGGGGTTTGCACCAATGAGCCCTGAAGATTATATTTG 924
Qy 793 AATACAAATGGAAGAGCTTTACTGATGATGAACTCGCGCAGCTCTATCTAGAAAAG 852
Db 925 ATCCCGATATG-----CAATGACCGATGAAGCAGCAGCCACTGTCCCGAAG 975
Qy 853 TCCCAATGCTTCTCCAAATAAATCCCTACAGAAATGGTCAATTTGTTCTGCGTGTGTTG 912
Db 976 TGCCAAATTCCTTCAAGCAAAATAAATCCATACCGAATGGTCAATTTGTAATTCGATGATAG 1035
Qy 913 TTCTAAGCATCTTCTCGACCTACGCTCTCACAAATCTGTGGTAAATGCAATACCCTGT 972
Db 1036 TGCTGGGTATTTTCTCCGCTATCGCTCTCTGGAATCCAGTGAAGATGCATATGGGCTCT 1095
Qy 973 GGCTTTTATCTGTTATATGTGAGATTTGGTTTGTGTTTATCTCTGGATCTGATCAGTTCC 1032
Db 1096 GGGCCACTTCTATCGTTTGTGAATCTGGTTTGGCTTGTCTGATGATTTCTTGTATCAGTTTC 1155
Qy 1033 CGAAGTGGTTTCAAATCAACCGGAGAGCTTACCTGTGATAGACTGGCTTTAAGGTATGACC 1092
Db 1156 CCAAGTGGTTGCCCTATCAGTCTGTGAAACGTATCTCTGATCGACTGTCTAATTAAGGTACGAAC 1215
Qy 1093 GAGAAGGTGAACCTCTCAGTTGCTGCTGTGACATATTTGTTCAGTACAGTTCGACCCCT 1152
Db 1216 GAGAAGGGCAACATCAATGCTTGACCTGTTGACCTTTTGTGAGTACTGTAGATCCAC 1275
Qy 1153 TGAAGGAGCCACCTATCGTCACTGCCAACACTGTGCTATCOATTTCTGCTGTGATTATC 1212
Db 1276 TGAAGGAGCCCTCTTGGTTACTGCCAATACAGTATTATCAATCCTTTTCAGTACTACTACC 1335
Qy 1213 CCGTGGACAAGGCTCTCTGCTATGATCTGATGAGGAGCTTCAATGCTGACTTTTGACG 1272
Db 1336 CTGTAGACAATGTGCTCTGTTATGCTCTGTATGCGGAGCGCTGATGCTTACTTTTGAAT 1395
Qy 1273 CATTTGGCTGAGACTTTCAGATTTCTAGGAATGGGTACCAATTTGTGAAGAAAGTATGACA 1332
Db 1396 CTCTCTGTGAGACCTCAGAAATTTGCCAGAAATGGGTACCAATTTGCAAGAAATTCGACA 1455
Qy 1333 TTGAACCCAGAGCTCCCGAGTTTACTTTTGGCCAGAAAATTTGATTACTGAAAGCAAAAG 1392
Db 1456 TTGAGCCTCGGCTCCCGAAATCTATTCTCTCAGAAAATTTGACTATCTGAAGGACAAAT 1515
Qy 1393 TCCAGCCTTCAATTTGTTAAAGACCGCGGCGCATGAGAGAGAAATGAAAGAAATTTAA 1452
Db 1516 TTCAACCCACCTTTGTCCNAGAGCGCGGTGCCATGAAGAGAGAAATATGAAGAAATTCAGG 1575

Qy 1453 TCAGGATAAATGCCCTAGTTTCTAAGGCATTGAAGTCCCGAGGAAGGATGGATCATGC 1512
Db 1576 TGCSCATCAATTCGGTTGGTTGCAAAAGGCTCTTAAAGTGCCCAAGGAAGGATGGCAATGC 1635
Qy 1513 AAGATGCAACACCATGGCCAGGAAAATAATACACAGGGATCATCTCTGGAATGATTCAGGTTT 1572
Db 1636 AAGAGGTACGCTTTGGCTTGGTAAATAATACCCGTGACCATCTCTGGTATGATCCAAGTGT 1695
Qy 1573 TCCTTGGTACAGTGTGGCTTGATPACTGAGGGTAATGAGCTCCCGGTTTAGTTTATG 1632
Db 1696 TCTTTGGTACAGTGGCGGCTCGATACAGAAGCAATGAGCTTCTCTCGGTAGTATATG 1755
Qy 1633 TGTCTCGTGAAGAGCGTCTGGGTTCCAGCACCAACAAGAGGCTGGTCCATGAATGCC 1692
Db 1756 TTTCTCGTGAAGAGAGACCTGGTTTCCAGCATACAAAGAGCCGCTGCCATGAATGCTT 1815
Qy 1693 TTGTTCTGTCTCAGCTGTCTTACTAATGGAACAATCATGTTGAAATCTTGAATGTGATC 1752
Db 1816 TGGTTCCGGTTTCTGCTGTGCTCACCAATGCTCCATTTATGCTGAATGTGATC 1875
Qy 1753 ACTACATCAACAACAGAGCTGTCCGAGAGCTATGTGCTTCTTCTAATGATCCAAACC 1812
Db 1876 ACTACATTAACAATAGCAAGGCAATCAGGGAAGGCATGTGCTTTATGATGGATCCTCAGG 1935
Qy 1813 TAGTCCGCAAGTCTGTTATGTGCAGTTTCCACAAAGGTTTGATGGGATTCATAGGAATG 1872
Db 1936 TTGGGAGAAAAGTCTGTTATGTCCAAATTCCTCAGAGATTCGATGGTATTCGCAATG 1995
Qy 1873 ATCGATATGCAAAACAGGAACACTGTCTTTTGTATATTAACCTTGAGGGGCTTCACGGCA 1932
Db 1996 ACCGTTACGCCAATCGAAACACCGTATTTCTTTGATATCAACATGAAAGGCTCTGGATGAA 2055
Qy 1933 TTCHAGACACAGTTTATGTGGAACTGTTGTGTTTTCACAGAACAGCTATCTATGGTT 1992
Db 2056 TTCAAGGGCTGTATATGTGGAACTGGATGTCATGTTTCAAGAACAAGCTCTATATGGGT 2115
Qy 1993 ATGAGCCCCCAATTAAGGGCGAAGAGCAGGTTTCTTGGGATCACTATGTGGGGCGAAGA 2052
Db 2116 ATGGGCTCTCCAAAGGCCCAAAAGCTCCCAAGATGGTGACTGTGATTGT-----C 2166
Qy 2053 AGAAGCAAGCAAGCTCAAAAGAAAGAGCTCAGATAAGAAAAGAGTCGAAACAAGCATGTGG 2112
Db 2167 TCCCTTGTGGGCTCTCGTAAGAAAGTCTCCGAAGAAAATAGTAGCAAGAAAAGTGCAG 2226
Qy 2113 ACAGTTCTGTTCAGTATTCATCTCGAAGACATAGAGGAGGTGTTTGAAGTCTGGGT 2172
Db 2227 GAATCCAGCTCCCGCTACAACTCGACCGGATCGAGGAAGGAGTGAAGGT-----T 2280
Qy 2173 TTGATGATGAGAAATCAGTTTCTCATGTCTCAAATGAGCTTAGAAGAGAGATTTGGCCAGT 2232
Db 2281 ATGATGACGAAGAGCATTTGTTGATGAGCCAACTAGACTTCGAGAAGAGTTTGGCCAGT 2340
Qy 2233 CAGCAGCATTTGTGTCCTCACTCTGATGGAATATGTTGGTGTTCCTCAGTCTCCCACTC 2292
Db 2341 CTTCAGCTTTTGTTCATCCACTCTGATGGAATGTTGGTGTTCCTCCGAAACAGCAAAATC 2400
Qy 2293 CAGAACTCTTTTGAAGAGAGCTATCCATGTCAATGTTGGCTATGAGGACAGTCTG 2352
Db 2401 CAGCTGAATTTGTTGAAGAGGCTTATCATGTCTCAGCTGTGGGATATGAAGACAAACGG 2460
Qy 2353 AATGGGAACTGAGATTGGTTGGATCTATGGATCTGTACAGAAAGATATTTCTACTGGAT 2412
Db 2461 AATGGGAAAGAGCTTTGGATGGATCTATGGATCAGTACAGAGACATTTCTGATGGAT 2520
Qy 2413 TCAAGATGCAAGAGGCTGGGCTTCAAGTCTAATGATGCCCAAGCCCGCAGCTTCA 2472
Db 2521 TCAAGATGCACTCGAGGCTGGGCTCCATTTACTGTATGCCCAAAACAGAGCAATTC 2580
Qy 2473 AGGATCTGCCCCCATCAATCTTTTTCAGATCTGTGAACCAAGTCTCGGTGGGCTCTCG 2532
Db 2581 AAGGCTCTGCTCCAATCAATCTATCAGACCGTTTGAACAGGTTTGGCTTGGCTTTGG 2640

```
QY 2533 GTTCTGTTGAAATCTTTTTCAGCGCGCATTTGCCCTTATGTTATGCTACGGAGGGCGCC 2592
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2641 GATCAGTAGAAATTTTCATGAGCAGACATTTGCCCAATCTGGTATGGCTATGGGGAGGTC 2700
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2593 TCAAGTTCCTGGAGAGATTCGTTTACATACACACCAATTTACCCATAACCTCTCTCC 2652
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2701 TGAATGGCTTGAAGATTTGCCCTATATCAACACCAATTTGCTATCCATTACCTCTCTTC 2760
QY 2653 CGCTTCTAGTCTATTTGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2712
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2761 CACTCATGCTATTTGACACACTTCCAGCCGTCAGTTTGCCTCACTGGCAAAATTTGTATCC 2820
QY 2713 CAGAGATTTAGCAACTTGGCCAGTATCTGGTTTCAATGCGCTCTCTTCAATTTTCGCCA 2772
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2821 CTGAGATCACTATTTTGGCAAGTCTATTTTATAGCTCTTTTCACTCTCAATTTTGGCA 2880
QY 2773 CTGCTATCTTGATGATGAGTGGAGTGGTGTGGCATGAGTGGTGGAGGAATGAAC 2832
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2881 CTGCTATCTTGAATGAGTGGAGTGGAGTGGATGAAGTGGTGGAGGAATGAAC 2940
QY 2833 AGTTCTGGGTCATTTGAGGATCTCTGACATCTGTTTGGCGTCTTTCAGGGTCTTCTGA 2892
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2941 AGTTCTGGGTCATTTGAGGAGTCTTCTGACATTTTTCAGTATTTTCAAGGTTCTGCTCA 3000
QY 2893 AGTGTCTGCGGTATCGACCACTTCACTGTCTCACTCAAAAGGCTAATGACGAAGAG 2952
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 3001 AGGTACTGGCAGGCAATGTATACAAATTTTCAAGTCACTGCGCAAGGC---ATCAGATGACG 3057
QY 2953 GCACATTTGCTGAGCTCTACATGTTTAAAGTGGAGCGCTTCTCATCCCTCGACGACCA 3012
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 3058 GTGAGTTTGGGGAACCTGTATGCAATTTCAATGGACCACTCTCTCAATCTCTTCTTACAACC 3117
QY 3013 TTTTGTATCATTAACATGTTGGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3072
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 3118 TGTGTGTCATCACTTGTGGGGTGGTGTGTTGGCTGAGATGCAATCAACATGGAT 3177
QY 3073 ACCAATCATGGGGCGCGCTTTTGGGAAGCTCTTCTTGGCTTCTGCTGCTGCTGCTGCTGCT 3132
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 3178 TTCAGTCACTGGGTCTCTCTTGGTAAAGTTTCTTGGATTTCTTGGTCAATTTGTGCACC 3237
QY 3133 TATACCATCTCTCAAGGTTCTTATGGGAGGGAACCAACCGCAACCGCAAGTGTCTATCG 3192
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 3238 TGTATCTTCTCTCAAGGTTCTCATGGGAGGAGGAACCAACCCACCATCGTGTGA 3297
QY 3193 TCTGGGCTGTCCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3252
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 3298 TTTGTCAATCTGCTGGCATCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3357
QY 3253 CTACCGCTCTGCTGGGCCCAATATCAAACTGTGGCATCAACTGCT 3300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 3358 TGAAGTAAAGTTAAAGGCCAGATCTAAACAATGTGGCATCAACTGCT 3405
```

```
RESULT 15
AAZ99530
ID AAZ99530 standard; DNA; 3753 BP.
XX
AC AAZ99530;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
DE Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key
FT 184..3405
FT /tag= a
FT /product= "cellulose synthase"
FT /trans_except= (pos: 916..918, aa: Xaa)
```

```
FT /note= "no termination codon given; Xaa is an unspecified amino acid"
FN WO200009706-A2.
XX 24-FEB-2000.
XX
XX 16-AUG-1999; 99WO-US018760.
XX 17-AUG-1998; 98US-0096822P.
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX WPI; 2000-224343/19.
XX P-PSDB; AAY84120.
XX
XX New genes which encode maize cellulose synthase polypeptides in plants
XX useful for modulating the expression of cellulose synthase in plants and
XX to produce transgenic plants expressing the novel protein.
XX
XX Claim 1; Page 184-188; 119pp; English.
XX
XX The present sequence encodes a maize cellulose synthase polypeptide. The
XX cellulose synthase can be used for the improvement of stalk quality for
XX improved stand or silage. It also provides an increased concentration of
XX cellulose in the pericarp, hardening the kernel and improving its
XX handling ability. The sequences are used to produce transgenic plants and
XX seeds expressing the cellulose synthase. The polynucleotide is used for
XX modulating, preferably increasing, the level of the synthase in a plant
XX cell. The plants are preferably monocots. The polynucleotide is also used
XX as a probe or primer in the detection quantitation or isolation of gene
XX transcripts. The probes are useful in detecting deficiencies in the level
XX of mRNA in screenings for desired transgenic plant, for detecting
XX mutations in the gene, for monitoring upregulation of expression or
XX changes in enzyme activity in screening assays of compounds, for use as
XX detection of any number of allelic variants of the gene, or for use as
XX molecular markers in plant breeding programs. The isolated nucleic acids
XX of the present invention can also be used for recombinant expression of
XX their encoded polypeptides or for use as immunogens in the preparation
XX and/or screening of antibodies. The proteins can be employed in assays
XX for enzyme agonists or antagonists of enzyme function or for use of
XX immunogens or antigens to obtain antibodies specifically immunoreactive
XX with a protein.
XX
XX Sequence 3753 BP; 944 A; 805 C; 965 G; 1038 T; 0 U; 1 Other;
XX
XX Query Match 37.9%; Score 1373; DB 3; Length 3753;
XX Best Local Similarity 66.3%; Pred. No. 0;
XX Matches 2130; Conservative 0; Mismatches 986; Indels 96; Gaps 7;
QY 53 CCGCGCCATGGACGGCGAGCGGCCCTCTGAAGTCCGGAGGACACGGGGCGGGGACGCT 112
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 113 GTGCCAGATCTGCGCGACGGCTGGGACACCGTTTGGACGGCAGCTCTTACCCGCTG 172
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 173 CGACGCTCTCCCGCTTCCCGCTGTCGGCCCTCTGCTACGAGCAGCAGCGGAGGCGAC 232
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 363 CAATGAGTGTGCTTCCCTCTGTCGGCCCAATGCTATGATGATGAGCGCAAGAGGGGAA 422
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 233 CCAGGCGCTCCCTCCAGTGAAGCAAGCAAGTACAGCGCCACAGAGGAGGCCAGCGATCG 292
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 423 CCAATCTGCCCCCAGTGAAGACTAGATAGAGAGACAGAAAGGTAGCCCTCGAGTTCA 482
QY 293 CGGGGAGGAGGAGCAGCAGCAGTGTGCCGATGATGATGATGATGATGATGATGATGATGAT 352
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 483 TGGTATGATGAGGATGAGGAAGATGTTGATGACCTAGACATGAATTAATCACTAC----- 534
QY 353 TGGCACTGAGGACCAAGACAGCAAGATTTGCTGACAGGATCGGAGCTGCGGATGACAC 412
```

Db 535 -----AACCAGGC 543
QY 413 CGGGGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTCTCCAA 472
Db 544 AGTGGGAAAGGCCCCAGAGTGGCACTGCAAG---GAGATGATGCTGCTCTTCATCT 600
QY 473 GTATGACAGTGGAGATCCCTAGGGATAGTCCCTTCAGTACCCACAGCAGATGTC 532
Db 601 GCTCGCCATGAGCCACACATCAATCGGATTCACGCGCTGACAAAGCGTCAA---CAGATATC 656
QY 533 AGGAGAAATCCCTCGAGCTTCGCTGATCATCATGATGATCCCTACGCGGGAACATCAG 592
Db 657 TGGAGAGATTCCTGATGCTTCCTGACCTGACCGTCACTCTATCCGAGTCCAAATCGAGCTA 716
QY 593 CAGACGTGCTCCGTTTCCCTATGGAATCATTCACCAAAATCCGTCACAGGGAGTTC---TC 649
Db 717 TGTGATCCAAAGCGTCCAGTTCCTGTGAGGATTTGGACCCCTCGAAGGACTTGAATTC 776
QY 650 CGGCAGTATGGGAATGTTGCTCGAAGAGAGAGATTTGATGGCTGGAAATGAAGCAGGA 709
Db 777 CTATGGGCTTAATAGTGTGATCTGGAAGAAAGAGTTGAGAGCTGGAGGTTAAACAGGA 836
QY 710 CAAGGGTGCATTCCTCATGACTAATGGGACCAAGCATTTGCTCCCTCTGAAGGTCCGGCAGC 769
Db 837 CAAAATATGATGCAAGTGACTAATAATATCCAGAGGCTAGAGGAGAGACATGGAGG 896
QY 770 TACTGACATCATGATCATCTGAAATCAACATGGAAGACGCTTTTACTGAATGATGAAC 829
Db 897 GACTGGCTCAATGGA-----GAANATATGCAAAATGGTTGATGTC 938
QY 830 TCGCCAGCTCTATCTAGAAAGTCCCATTTGCTTCCTCCAAATATAATCCCTACAGAT 889
Db 939 ACGGCTACCTTTGAGCCGTATCGTGCCAAATTTCTTCAACACAGCTCAACCTTTACCGGT 998
QY 890 GGTCAATTTCTCGGGTGGTGTGTTTAAGCATCTTCTGCACTACCGTCTCACAATCC 949
Db 999 AGTGATCATTTCTCGTCTTATCATCTCTGCTCTTCTTCCAGTATCGTGTGAGTCACTC 1058
QY 950 TGTGGTAAATGATACCACTGTCGCTTTTATCTGTTATATGATGAGATTTGGTTGCTTT 1009
Db 1059 AGTGGTGATGCTTATGATGATGCTAGTATGCTGTTATCTGCGAGGTCTGGTTTGCCTT 1118
QY 1010 ATCTGGATATCTGATCAGTATCCGAGTGTTCCTCAATCAACCGGAGACTACTCTTGA 1069
Db 1119 GTCTTGGCTCTAGATCATGTTCCCAAATGGTATCCAAATCAACCGTGAGACATCTTGA 1178
QY 1070 TAGACTGGCTTTAAGGTATGACCGAGAGGTGAACCGTCTCAGTTGGCTGCTGTTGACAT 1129
Db 1179 CAGGCTTGCATTTGAGTATGATAGAGAGGAGAGCCATCACAGCTGGCTCCCATTTGATG 1238
QY 1130 ATTTGTGATGACGTCGACCCCTTGAAGAGGACCACTATCGTCACTGCCAACACTGTGCT 1189
Db 1239 CTTCTGATGACGTTGATCCATTTGAAGGAACTCCCACTGATCACAGCCAACTGTTTT 1298
QY 1190 ATCCATTTCTGCTGTGATATTCGCTGACAAAGTCTCTTCTGCTATGTTATCTGATGACG 1249
Db 1299 GTCCATTTCTTCTGTTGATTTACCTGTTGACAAAGTGTCACTGATGTTCTGATGATGG 1358
QY 1250 AGCTTCAATGCTGACTTTTGAAGCATGCTGCTGAGACTTCAGAGTTTGTAGGAAATGGT 1309
Db 1359 TTCAGCTATGCTGACTTTTGTAGTCTCTCTGAAACCGCAGAAATTTGCTAGAAAGTGGT 1418
QY 1310 ACCATTTGTGAGAGATGATGATTTGAACCCAGAGCTCCGAGTTTACTTTTGCAGNA 1369
Db 1419 TCCCTTTTGTAAAGAGCAATATTTGAACCAAGAGCTCCAGATTTTACTTTTGCCTCAAA 1478
QY 1370 AATTGATTAACCTGAAAGCAAGTCCAGCTTTCAATTTGTTAAAGACCGCGCGGCATGAA 1429
Db 1479 AATAGATTACCTGAGGACAAATTTCACTTTCAATTTGTTAAAGAAAGACGCGCAATGAA 1538
QY 1430 GAGAGATATGAGAAATTTAAATCAGGATAAATGCCCTAGTTTCTTAAGGCTATTGAAGT 1489

1539 GAGGAGTATGACGAATTCAAAGTAAAGATCAATGCCCTTGTGTCCAAAGCACAGAAAGT 1598
QY 1490 CCCCGAGGAGGATGATCATGCAAGATGACACACCATGCGCAGGAAACAATAACAGGGA 1549
Db 1599 GCTTGAAGAGAGGAGTGGACCATGCTGATGGAATGCAATGGCTGGGATTAATCTTAGGGA 1658
QY 1550 TCATCTCTGGAATGATTCAGGTTTTCTTGGTTCACAGTGGTGGCTTGTACTCTGAGGTTAA 1609
Db 1659 CCATCTGGCATGATTCAGGTTTTCTTGGGGCACAGTGGTGGCTCGACACTGATGAAA 1718
QY 1610 TGAGCTCCCCGTTTATGTTATGTTCTCGTGAAGAGGCTCTGGGTTTCAGACACCAAA 1669
Db 1719 TGAGTTACCACTCTTCTGTTATGTTCTCGTGAAGAGACAGGCTTTTCAGATCAAA 1778
QY 1670 GAAGGCTGGTCCCATGATGCCCTTGTTCGTCTCAGCTGCTCTTACTTAATGACAATA 1729
Db 1779 GAAGGCTGGTGAATGAATGCGCTGATTCGTGTATCTGCTGTCTGACAAATGGTGCCTA 1838
QY 1730 CATGTTGAATCTTGATTTGATCTACATCAACAACAGCAAGGCTGTCCGAGAAGCTAT 1789
Db 1839 TCTTCTCAATGTGATTTGGACCATTTACTTCAATAGCAGCAAGCTCTTAGAGAAGCAAT 1898
QY 1790 GTGCTTCTTAATGATCCAAACCTAGTCCGCAAGTCTGTTATGTCAGTTTCCACACAAG 1849
Db 1899 GTGCTTCTATGATGATCGGCTCTAGGAAGGAAAACTTTGTTATGTAACAATTTCCACAG 1958
QY 1850 GTTTGATGGGATTCATAGGAATGATGATATGCAAAACAGGAACACTGTCTTTTGTGATAT 1909
Db 1959 ATTTGATGGCATGACTTTCAGATGATGATGCTTAATCGGAACATAGTTTCTTTGATAT 2018
QY 1910 TAACTTGAGGGGCTTGAACGCAATCAAGGACCAAGTTTATGTTGGAACCTGTTGTGTTTT 1969
Db 2019 CAACATGAAGGTCTGGATGSCATTCAGGCTCCAGTTTACGTGGGAACAGGATGCTGTT 2078
QY 1970 CAACAGAACAGCTATCTATGTTATGAGCC---CCCAATTAAGGCGAAGAACCCAGGTTT 2026
Db 2079 CAATAGACAGGCTTTGTATGATGATGATCTCTGTTTGAAGCTGATCTGAGGCCAAA 2138
QY 2027 CTTGCGCATCACTATGTGGGGCAAGGAAGCAAGCAAGTCAAGAAAAAGAGCTCAGA 2086
Db 2139 CATTTGTTATTAAGAGCTGCTGTGTGATGAGGAAGAAAGAACAGAGTTATATGATAG 2198
QY 2087 TAAGAAAAAGTGAACAGCATGTGGACAGTCTTGTTCAGATTTCAATCTCGAAGACAT 2146
Db 2199 TCAAGCGGTATTATGAAGAACAGAAATCTTCAGCTCCCATCTTCAATATGAAGACAT 2258
QY 2147 AGAGAGGCTGTTGAAGGTGCTGGTTGATGATGAGAAATCAGTTCTCATGCTCAAT 2206
Db 2259 CGAAGAGGTTATTGAAGT-----TACGAGGATGAAGGTCAGTGTATGTCACAG 2312
QY 2207 GAGCTTAGAGAGAGATTTGGCCAGTACAGAGCATTTTGTGCTCCACTCTGATGGAATA 2266
Db 2313 GAAATGGAGAAACGCTTTGCTGCTCTCTTATTTTCAITGATCCACTTTATGACACA 2372
QY 2267 TGTGTTGTTTCTCCTCAGTCCCTCCACTCCAGAAATCTCTTTTGAAGAGCTATCCATGCT 2326
Db 2373 AGTGGGCATACCACTTCAACAAACCCAGCTTCTTACTATAAGGAAGCTATCCATGTCT 2432
QY 2327 AAGTTGTGCTTATGAGGACAGTCTGATGAGGAACTGAGATTTGTTGATCTATGATC 2386
Db 2433 CAGTTGTGATGATGAGGACAAACTGAATGGGAAAAGAGATTTGGCTGATCTATGTTTC 2492
QY 2387 TGTACAGAAATATTTCTTACTGGAATTCAGATGCAACGACAGGCTCGGCTTCACTCTA 2446
Db 2493 AGTAACGAGGATATTTGACTGGTTTAAATGATGATGCAAGGGGCTGGCAATCAATCTA 2552
QY 2447 TTGATGCCCAAGGCCCGAGCTTTCAAGGATCTGCCCCCATCAATCTTTGATGCTGCT 2506
Db 2553 CTGATGCCACCAACGACCTTTGTTCAAGGTTTCTGACCAATCAATCTTTCCGATCGCT 2612
QY 2507 GAACCAAGTGTGCGTGGCTCTCGTTTCTGTTGAAATCTTTTTCAGCCGCAATGCC 2566
Db 2613 TAATCAGGTGCTCCGTTGGGCTCTTGGGTGAGTGAATCTGCTTAGTACATGTTGCC 2672

QY	2567	CTTATGGTATGCTACGAGGCGGCTCAGTTCCTGGAGAGATTGCGTTACATCAACAC	2626
Db	2673	TATCTGGTATGGTTACAATGGACGATTGAAGCTTTTGGAGAGGCTGGCTTACATCAACAC	2732
QY	2627	CACCAATTTACCCACTAACTCTCTCCGCTTCTAGTCTATTGTATATTGCTTGCCTGCTATCTG	2686
Db	2733	TATTGTATATCCAAATCAGATCCATTCGCTTATTGCCTATTGTGTCTCCGCTATCTG	2792
QY	2687	TCCTGCTACTGGAAAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTTCAT	2746
Db	2793	CTCTCCTTACCAATAAATTTATCATCTCTGAGATTAGCAATTATGCTGGGATGTTCTTCAT	2852
QY	2747	TCGGCTCTTCTCTTCAATTTTCCCACTGGTATCCTTGGAGATGAGGTGAGTGGTGTGG	2806
Db	2853	TCCTCTTTTTCGGCTCCATTTTTCGCACTGGTATATTGGAGCTTAGATGGAGTGGTGTGG	2912
QY	2807	CATTGACGAGTGGTGAGGAATGAACAGTCTCTGGGTCAATTGGAGGTATCTCTGCACATCT	2866
Db	2913	CATTGAGATTGGTGAGGAATGAGCAGTTTGGGTATTGGTGGCACCTCTGCCCACT	2972
QY	2867	GTTCGCGTCTTTACGGTCTCTGAGGTGCTTGCCTGATCGACACCAACTTCACTGT	2926
Db	2973	CTTCGAGTGTTCACGGGTCTGTGAAAGTGTGGCTGGGATTGATACCAACTTCACAGT	3032
QY	2927	CACCTCAAGGCTAATGACGAGAGGAGGACCTTCTGCTGAGCTTACATGTTCAAGTGGAC	2986
Db	3033	TACCTCAAGGCTAATGATGAGATGGGACCTTCTGCTGAGCTATATGTTTCAAGTGGAC	3092
QY	2987	GACGCTTCTCATCCCTCCGACGACCAATTTTGATCATTAACATGGTGTGTGTTGCTGG	3046
Db	3093	CAGTTTGTCTCATCTCTCCGACCACTGTTCTTGTCTATTAACTGGTGGGATGGTGGCAGG	3152
QY	3047	CACCTCTACGCCATCAACAGTGGTTTACCAATCATGGGGGCGCTCTTTTGGGAAGCTCTT	3106
Db	3153	AATTTCTTATGCCATTAAACAGTGGCTACCAATCCTGGGTCCGCTCTTTTGGAAAGCTGT	3212
QY	3107	CTTTGCGCTTCTGGGTGATTGTTCACTTATACCAATTCCTCAAGGGTCTTATGGGACGCA	3166
Db	3213	CTTCTCGATCTGGGTGATCCTCCATCTCTACCCCTTCTCAAGGGTCTCATGGGAAGGCA	3272
QY	3167	AAACGCGACACGAGATTGTCATCGTCTGGGCTGTCCTCTCGCTTCTATCTTCTCCTT	3226
Db	3273	GAACCGCACACCAACAAATCGTCAATGCTGGTCCATCCCTTCTTGCACTATCTTCTCCTT	3332
QY	3227	GCTGTGGGTTCGTGTTGATCCATTCATCCC	3258
Db	3333	GCTGTGGGTGAAGATCGATCCCTTTCATCTCCC	3364

Search completed: August 23, 2004, 01:39:31
Job time : 1325 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2004, 01:15:00 ; Search time 236 Seconds
(without alignments)
8526.491 Million cell updates/sec

Title: US-09-900-237A-29
Perfect score: 3626
Sequence: 1 gcacagaggaacccgctcca.....tcacattttggaggagtttt 3626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1673.6	46.2	3614	4	US-09-221-013A-9 Sequence 9, Appli
2	1343.4	37.0	3603	4	US-09-221-013A-5 Sequence 5, Appli
3	1341.8	37.0	3673	4	US-09-221-013A-11 Sequence 11, Appli
4	1215.6	33.5	3328	3	US-08-960-048-1 Sequence 1, Appli
5	1215.6	33.5	3328	4	US-09-838-586-1 Sequence 1, Appli
6	1176.6	32.4	3828	4	US-09-221-013A-7 Sequence 7, Appli
7	984.2	27.1	2248	4	US-09-221-013A-1 Sequence 1, Appli
8	568.8	15.7	1741	4	US-09-221-013A-13 Sequence 13, Appli
9	491.6	13.6	8411	4	US-09-221-013A-3 Sequence 3, Appli
10	219.6	6.1	5009	4	US-09-221-013A-4 Sequence 4, Appli
11	206.4	5.7	281	4	US-09-313-294A-6485 Sequence 6485, Ap
12	204	5.6	300	4	US-09-313-294A-4753 Sequence 4753, Ap
13	198.2	5.5	291	4	US-09-313-294A-5519 Sequence 5519, Ap
14	177	4.9	286	4	US-09-313-294A-4696 Sequence 4696, Ap
15	146.8	4.0	284	4	US-09-313-294A-3993 Sequence 3993, Ap
16	140.2	3.9	268	4	US-09-313-294A-849 Sequence 849, App
17	111.6	3.1	278	4	US-09-313-294A-4015 Sequence 4015, Ap
18	104.8	2.9	159	4	US-09-313-294A-1693 Sequence 1693, Ap
19	102.6	2.8	287	4	US-09-313-294A-5274 Sequence 5274, Ap
20	96.8	2.7	283	4	US-09-313-294A-1981 Sequence 1981, Ap
21	94.6	2.6	317	4	US-09-313-294A-7123 Sequence 7123, Ap
22	81.2	2.2	250	4	US-09-313-294A-582 Sequence 582, App
23	59	1.6	3624	1	US-07-951-715A-6 Sequence 6, Appli
24	59	1.6	3624	2	US-08-459-448A-6 Sequence 6, Appli
25	59	1.6	3624	3	US-08-459-595A-6 Sequence 6, Appli
26	59	1.6	3624	3	US-08-459-504B-6 Sequence 6, Appli
27	59	1.6	3624	3	US-08-459-444-6 Sequence 6, Appli

28 59 1.6 3624 3 US-09-053-549-7 Sequence 7, Appli
29 59 1.6 3624 4 US-09-547-422-6 Sequence 6, Appli
30 55.4 1.5 390 3 US-09-197-649-7 Sequence 7, Appli
31 55 1.5 7218 1 US-08-232-463-14 Sequence 14, Appli
32 53.4 1.5 28958 1 US-08-258-261B-6 Sequence 6, Appli
33 53.4 1.5 28958 1 US-08-456-837-6 Sequence 6, Appli
34 53.4 1.5 28958 1 US-08-457-342-6 Sequence 6, Appli
35 53.4 1.5 28958 1 US-08-457-646A-6 Sequence 6, Appli
36 53.4 1.5 28958 1 US-08-458-076A-6 Sequence 6, Appli
37 53.4 1.5 28958 1 US-08-764-233A-4 Sequence 6, Appli
38 53.4 1.5 28958 1 US-08-457-335A-6 Sequence 6, Appli
39 53.4 1.5 28958 1 US-08-729-214-6 Sequence 6, Appli
40 53.4 1.5 28958 3 US-09-028-934-6 Sequence 6, Appli
41 53.4 1.5 49377 1 US-08-764-233A-1 Sequence 1, Appli
42 51.4 1.4 4403765 3 US-09-103-840A-2 Sequence 2, Appli
43 49.2 1.4 4411529 3 US-09-103-840A-1 Sequence 1, Appli
44 48.8 1.3 4612 3 US-08-960-048-2 Sequence 2, Appli
45 48.8 1.3 4612 4 US-09-838-586-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-221-013A-9
; Sequence 9, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3614
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (217)..(3411)
US-09-221-013A-9

Query Match 46.2%; Score 1673.6; DB 4; Length 3614;
Best Local Similarity 71.4%; Pred. No. 0;
Matches 2347; Conservative 0; Mismatches 869; Indels 72; Gaps 8;

QY 30 TCGTCCGGTGGATCGCTCTCCGCGCCATGAGCGGCGGACCGCCCTGAGTCC 89
Db 190 TCGTGTGGAAGCACTAAGTGACAAATCGAAGGAGAAACCGCGGAAGCG 249
QY 90 GGAGGACCGGGCGGCGCGTGTGCCAGATCTGGCGCGCGCGCTGGGCACCGTTG 149
Db 250 ATGAAGACATGTTCCGCGAGACTTGCAGATCTGTAGTGACAAATTTGGCAGACTGTT 309
QY 150 GACGGCGAGTCTTACACCGCTCGGACGCTCGCGCTTCCCGTTCGCGCCCTGCTAC 209
Db 310 GATGGAGATCGTTTGTGCTTGTGATATTTGTTCATTCACAGTTTGTGCGCTTCTAC 369
QY 210 GAGCAGCGGCAAGGAGGCGCCAGGCTGCTCCAGTGCAGAGCAAGTACAGCGC 269
Db 370 GAGTATGAGAGGAAGATGGGAATCAATTTGTCTCAGTGCAAAACCAAGATCAAGAG 429
QY 270 CACAGAGGAGCGCCAGCGATCCGCGGGAGGAGGCGACACTGATCCGATGATGTT 329

2478 TCTGCCCCCAATCAATCTTTCAGATCGTCAACCAAGTCTCGGTGGCTCTCGTTCT 2537
2590 TCTGCTCCATCAATCTTTCAGATCGTCAACCAAGTCTCGGTGGCTCTCGTTCTCA 2649
2538 GTTGAATTTCTTTTCAGCGGCAATGCCCTTATGGTATGGTACGAGGGCGCTCAAG 2597
2650 GTTGAGATTTCTTTCAGTGGCAATGGCTTATATGGTATGGTACAAATGGAGGCTAAAA 2709
2598 TTCTCTGGAGATTCGTTACATCAACACCAATTTACCCACTAGCTCTCTCCGCTT 2657
2710 TTCTTTGAGAGGTTTGGTATGTAACACCAATCTACCCCTATCACTTCTCTCTCT 2769
2658 CTAGTCTATTGATATTTGCTGCTATCTGCTCACTGGAAGTTCAATGACCAAG 2717
2770 CTCAATGTTTACATTTAGCGGTTTGTCTTCAACACAGTTTATATCTCTAG 2829
2718 ATTAGCAACTTGGCCAGTATCTGGTTCAATGCGCTCTCTCTTCAATTTTGGCCACTGGT 2777
2830 ATTAGTAACATTTGCAAGTATATGGTTTCTCTCTCTCTCTCTCTCTCTCTCTCT 2889
2778 ATCTTTGAGATGAGTGGAGTGTGTGGATTTGAGATTTGAGAGTGGAGGAATGAACAGTTT 2837
2890 ATACTAGAAATGAGTGGAGTGGCTAGGATAGACCAATGGTGGAGAAACGAGAGTTT 2949
2838 TGGGTCAATTTGGAGGTATCTCTGACATCTGTTTGGCGTCTTTTCAGGTCTTTCTGAAGGTG 2897
2950 TGGGTCAATTTGGAGGTATCTCTGATTTATTTGCTGTGTTTCAAGGTATCTTCAAGTTC 3009
2898 CTTCGCGGTATCGACCAACTTTCACTGTCACCTCAAGGCTTATGACGAAGAAGCGAC 2957
3010 CTTCGCGGTATCGACCAACTTTCACTGTTTACCTCAAGGCTTCAATGAAGACGGAGAC 3069
2958 TTGCTGAGCTCTACATGTTCAAGTGAACAGCGCTTCTATCCCTCGGAGACCATTTTG 3017
3070 TTGCTGAGCTCTACTTGTTCAAATGGACAACTTCTGATTCGCGCAACGAGCTGCTC 3129
3018 ATCATTAACATGTTGGTGTGCTTGGCGACCTCTTACGCCATCAACAGTGTGTTACCAA 3077
3130 ATTGTAACCTTAGTGGAGTGTGAGAGTCTTATGCTATCAACAGTGGATACCAA 3189
3078 TCATGGGGCGCTCTTTGGGAAGCTCTCTTTGGCTTCTGGGTGATTTGTTCACTTATAC 3137
3190 TCATGGGGACCACTCTTTGGTAAGTGTCTTTGGCTTCTGGGTGATTTGTTCACTTGTAC 3249
3138 CATTTCTCAAGGTCTTATGGCAGCAAAACCGACACCGAGGATTTGCTATGCTGTG 3197
3250 CCTTCTCTCAAGGTTTGTGGTTCGACAGAACCGACTCTTACCATTTGTTGGTCTGG 3309
3198 GCTGTCTCTCTCTATCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3257
3310 TCTGTTCTCTGCTTCTATCTTCTGTTGTTGGGTAGGATTTGATCCCTTCACTAGC 3369
3258 GGTCTCGCTGGCCCAATATCCAAACCTTGGGATCACTGCTAGGAA 3305
3370 CGAGTCACTGGCCCGGACATTTCTGGAATGTGGAATCAACTGTTGAGAA 3417

RESULT 2

US-09-221-013A-5
; Sequence 5, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699

; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 3603
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3243)
US-09-221-013A-5

Query Match

37.0%; Score 1343.4; DB 4; Length 3603;
Best Local Similarity 66.2%; Pred. No. 0;
Matches 2098; Conservative 0; Mismatches 991; Indels 80; Gaps 8;

QY 92 GAGGCACGGGGCGGGAGCTGTGCCAGATCTGCGCGCAGCGGCTGGGCACCACTGTGA 151
Db ||||| GAAGATATGAATGGCCAGATATGTCAGATCTGTGTGATGATGTTGGACTCGCTGAAC 152
QY 152 CGCGCAGCTCTTCAACCGCTGCGAGCTCTGCGGCTTCCCGCTCTGCGGCTCGGCGCTGCTACGA 211
Db ||||| TGGAGATGTTTGTGCGGCTGTAATGAATGTGCTTCCCTGTGTGCTGCGGCTTGTCTATGA 212
QY 212 GCACGAGCGCAAGGAGCGGACCCAGCGCTGCCCTCCAGTCCAGACCAAGTACCAAGCGCCA 271
Db ||||| GTACGAGAGAAAGATGGAACCTCAGTGTGCCCTCAATGCAAGCTAGATTCAAGACACA 272
QY 272 CAGAGGGAGCCAGCGATCCGCGGGAGGAAGCGACGACCTGATGCCGATGATGTTAG 331
Db ||||| CAGGGGAGTCTCGTGTGTAAGAGATGAAGATGAGGATGATGTTGATGATATCGAGAA 332
QY 332 TGACTTCACTACCTCTGCAATCTGGCACTGAGGACCAAGAGAGATGCTGACAGGAT 391
Db ||||| TGAGTTCAATTACGCT 365
QY 392 GCGCAGCTGCGCGATGAACACCGGGGAGTGGCAATGTTGGCCACCCCAAGTATGACAG 451
Db ||||| GAGACACCAACCGCTTGGCGAAGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 425
QY 452 TGGCGAGATCGGCTCTCCAAAGTATGACAGTGGAGAGATCCCTTAGGGGATACGCTCTTC 511
Db ||||| TCTCTCTCTCAACCATGGCCATACGCTTCTGGAGAGATGCGACGCTGATACACAATC 485
QY 512 AGTCAACCAAG-CGAGATGTGAGGAGAAATCCCTGGAGTTCCGCTGATCATCATGA 570
Db ||||| TGTGCGAACTACATCAGGTCCTTTGGGTCCTTCTGACAGGAATGCTATTCTCTCCATA 545
QY 571 TGTCCCTACGGGGAACATCAGCAGCGTCTCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 630
Db ||||| TATGATCCACGGCAACCTGCT 599
QY 631 ATCCGTCAGGGAGTTCTCCGCGAGTATTGGGAATGTTGGGAATGTTGCTGGAAGAGAGATTGAT 690
Db ||||| GAACTCTTATGGGCTTGGTAATGTTGACTGGAAGAAAGAGATTGAAG 646
QY 691 GCTGGAATGAGAGAGGACCAAGGTGGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 750
Db ||||| GCTGGAAGCTGAGCAGGAGAAATATGTTACAGATGAC- - - - - TGGTA 691
QY 751 CTCTCTGAAGTGGGCGAGCTACTGACATCGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 810
Db ||||| AATACCATGAAGGAAAGGAGAGAAAT- - - - - TGAAGGAGTGGTTCATGCGGAAGAAC 748
QY 811 CTTTACTGAATGAATACTCGCCAGCT 870
Db ||||| TCCAAATGCTGATGATACAGCT 808
QY 871 AATAAATCCCTACAGAAATGGTCAATTTCTGCGGTTGGTGTGTTGTTGTTGTTGTTGTTGTTGTT 930
Db ||||| GCCTAACCCCTTATCGGGTTGATTTCTCGGCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 868
QY 931 ACTACGCTCTCACAATCTCTGTCGCTAATGACATACCCACTGTCGCTTTTATCTGTTATAT 990

Db 869 AATATCGTACAACTCACCCCTGTGAAATGATATCTTTGGTGTGACCTCGGTATCT 928
Qy 991 GTGAGATTTGGTTTGTCTTATCTCTGGAATCTGGATCTCCGAAAGTGGTTTCAATCA 1050
Db 929 GTGAGATCTGGTTTGCATTTTCTTGGCTTCTTGATCAGTTTCCCAATGGTACCCCAATTA 988
Qy 1051 ACCGGGAGACCTACTCTGATAGACTGGCTTTTAAGGTATGACCGAGAAGTGAACCGTCTC 1110
Db 989 ACAGGGAGACTTATCTTGACCGTCTCGCTATAAGATATGATCGAGACGGTGAACCATCAC 1048
Qy 1111 AGTTGGCTGCTGTTGACATATTTGTCAGTACAGTCGACCCCTTGAAGGAGCCACCTATCG 1170
Db 1049 AGCTCGTTCCTGTTGATGCTGTTGTTAGTACAGTGGACCCCAATTGAAAGAGCTCCCTTG 1108
Qy 1171 TCATGCGCAACACTGTGCTATCCATCTTTGCTGTGTGATTAATCCCGTGGACAAAGTCTCT 1230
Db 1109 TTACAGCAAAACACAGTTCTCTCGATCTCTCTGATCTCTGTTGGACTACCCGGTAGATAAAGTAGCCT 1168
Qy 1231 GCTATGATCTGATGACGAGCTTCAATGCTGACTTTTGAAGCTATGACGCTATGGCTGAGACTTCAG 1290
Db 1169 GTTATGTTTCAGATGATGGTTCAGCTATGCTTACCTTTTGAATCCCTTTCTGAACCGCTG 1228
Qy 1291 AGTTTGTAGGAATGGGTACCAATTTGTGAAGAAGTATGACATTTGAACCCAGAGCTCCCG 1350
Db 1229 AGTTTGAAGAAGTGGTACCAATTTTGAAGAAGTATGACATTTGAACCCAGAGCTCCCG 1288
Qy 1351 AGTTTACTTTTGGCCAGAAAATTTGATTAATCTGTAAGACAAAGTCCAGCTTCAATTTGTTA 1410
Db 1289 AATTCTAATTTTGGCCAGAAAGTAGATTAATCTTGAAGGACAAAGTCCAAACCGTCTTTTGTGA 1348
Qy 1411 AAGACCGCGGCGCATGAAGAGAGATATGAAGATTAAGATTTAAATCAGGATTAATCCCTAG 1470
Db 1349 AAGAGCGACGAGCTATGAAGAGAGAGATGAAGATTTAAAGTGAAGATTAATGCTCTTG 1408
Qy 1471 TTTCTTAAGGCAATGAAAGTCCCGAGGAGGATGATCATGCAAGATGGCACACCATGGC 1530
Db 1409 TTGCAAGACACAGAAATCCCTGAAGAGGCTGGACAAATGCAAGATGCTACCTCGC 1468
Qy 1531 CAGGAACAATACAGGAGATCATCTCGAATGATTTCAAGTTCCTTGGTCAAGTGTG 1590
Db 1469 CTGATAACACACTAGAGATCATCTCGAATGATACAGGTGTTCTTAGGCCATATGTGGG 1528
Qy 1591 GCCTGATACAGGATTAAGCTCCCGTTAGTTTATGTTTATGTTCTCGTGAAGAGCTC 1650
Db 1529 GTCTGGATACCGATGGAATGAGTGCCTAGACTCATCTATGTTTCTCGTGAAGAGCGC 1588
Qy 1651 CTGGTTCACAGCACCAAGAGGCTGGTGGCATGAATGCCCTTGTCTGCTCAGCTG 1710
Db 1589 CTGGATTTCAACACCAACCAAGAGGCTGGAGCTATGAATGCAATCGTGTATCTGCTG 1648
Qy 1711 TCCTTACTAATGGACAAATACATGTTGAATCTTGAATGTTGATCACTACATCAACACAGCA 1770
Db 1649 TTTCTACCAATTTGGAGCATATCTTTTGAACGTTGATGATCACTTACTTAAATACAGTA 1708
Qy 1771 AGGCTGTCGAGAGACTATGCTTCTTAAATGATTCACAACTAGTCCGCAAGTCTGTT 1830
Db 1709 AGGCTAATTAAGAGCTATGTTTCAATGATGGACCCCGCTATTTGGAAGAAGTCTGCT 1768
Qy 1831 ATGTGCACTTCCCAAAAGGTTTGAATGGATTTGATAGGAATGATCGATATGCAACAGGA 1890
Db 1769 ATGTCCAGTTCCTCAACGTTTGGACGGTATTTGATTTGACAGTATGATGATGCAACAGGA 1828
Qy 1891 ACATGCTCTTTTGGATTAATCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1950
Db 1829 ATAATGCTTTTGGATTAATCAATGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1888
Qy 1951 TGGAACTGTTGTTTTCACAGAAACAGCTATCTATGTTTATGAGCC-----CCCA 2003
Db 1889 TGGTACTGTTGTTTAAATAGGAGGCTCTATATGGGTATGATCTGTTTGTGACGG 1948
Qy 2004 ATTAAGGCAAGAACCCAGGTTTCTTGGCATCA--CTATGTGGGGGCAAGAGAGGCA 2061

Db 1949 AAGAAGATTTAGAACCAAAATATTATTGTCAAGAGCTGTTGCGGTCAAGGAAGAAGGTA 2008
Qy 2062 GCAAGTCAAGAAAAGAGCTCAGATAAGAAAAGTCGAAACAAGCATGTGACAGTTCTG 2121
Db 2009 AAAGTAGCAAGAAGTATAACTACGAAAGAGAGAGGAGCATCAACAGAAAGTCACTCCAATG 2068
Qy 2122 TTCCAGTATTCAATCTCGAAGACATAGAGAGGCTGTTGAAGGTGCTGCGGTTTGAATGATG 2181
Db 2069 CTCACATTTTCAATATGAGGACATCGATGAGGCTTTTGAAGGT-----TATGATGATG 2122
Qy 2182 AGAATCAGTTTCTCATGCTCTCAATGAGCTTTAGAGAAGAGATTTGGCCAGTCAGCAGCAT 2241
Db 2123 AGAGTCTATTTCTAATGTCGCCAGAGGAGTGTAGAGAAGCGTTTGGTCAAGTCCGCGTAT 2182
Qy 2242 TTGTTGCCCTCACTCTCATGATGAATATGTTGTTTCTCAGTCTCTCACTCTCAGATCTC 2301
Db 2183 TTATTTGGGCAACCTTCAATGAACAAGCGGCAATTTCCACCAACAACCAATCCCGTACTC 2242
Qy 2302 TTTTGAAGAAGAGCTATCATGTCATAGTTCTGCTATGAGCAAGTCTCAATGGGAA 2361
Db 2243 TTCTGAAGAGGCTATTTCAATGTTATAGCTGTGTTACGAAGAACAAGTCTGAATGGGCA 2302
Qy 2362 CTGAGATTTGGTGGATCTATGATCTGTCAAGAGATATTTCTTACTGATTTCAAGATGC 2421
Db 2303 AAGAGATTTGGTGGATCTATGTTCCGTGACGGAAGATATTTCTTACTGGTTCGAATGC 2362
Qy 2422 ACGCAAGAGGCTGGGTTCACTTATGTCATGCCCAGCGCCAGCTTTCAAGGATCTG 2481
Db 2363 ATGCCCGGGTGGATATCGATCTACTGCAATCTCCACGCCCTGCGTTCAAGGAGCTG 2422
Qy 2482 CCCCCTCAATCTTTTCAGATCGTCTGAACCAAGTCTGCGGTGGGCTCTCGGTTCTGTTG 2541
Db 2423 CACCAATCAATCTTTCTGATCGTTGAAACCAAGTCTTTCGATGGGCTTTGGGATCTATCG 2482
Qy 2542 AAATCTTTTTCAGCGGCAATTTGCCCTTATGTTATGCTACGAGGCGGCTCTCAAGTTC 2601
Db 2483 AGATTTCTTTAGCAGACATTTGCTCTATCTGATGTTTACCATGGAAGTGTGAGACTTT 2542
Qy 2602 TGGAGAGATTTGGCTTATCATCAACACCACTTATCCCACTAACCTCTCTCCGCTCTAG 2661
Db 2543 TGGAGAGATCGCTTATATCAACACCACTGCTCTATCTTATTCATCCATCTCTTATG 2602
Qy 2662 TCTATTGTATTTGCTGCTATCTGCTCTGCTCACTGGAAGTTCATCATGCGCAGAGATTA 2721
Db 2603 CGTATTGTATTTCTCCGCTTTTGTCTCATCAACGACAGATTCATCATATCCGAGATAA 2662
Qy 2722 GCAATTTGGCAGTATCTGTTCAATGCGCTCTTCTCTTCAATTTTCCGCTCTGATATCC 2781
Db 2663 GCAACTACGCGAGTATTTGGTTCTATCTTCTCTCTCATCTCAATTTGCTGTGACTGGAATCC 2722
Qy 2782 TTGAGATGAGTGGAGTGGTGTGGCATTTGAGAGTGGTGGAGGAAACAGTCTCTGG 2841
Db 2723 TGGAGCTGAGATGGAGCGGTGTGAGCATTTGAGGATTTGGTGGAGAACGAGCAGTTCTCGG 2782
Qy 2842 TCATTTGAGGATCTCTGTCATCTGTTTGGCTCTTTTCAAGGCTCTTCTGAAAGTGTCTG 2901
Db 2783 TCATTTGGTGGCAGATCCGCCATCTTTTGTCTCTTCCAAAGTCTACTTAAAGTGTCTG 2842
Qy 2902 CCGGTATGCAACCACTTCACTGTACCTCAAGGCTTAATGACGAAGAAGCGGACTTTG 2961
Db 2843 CTGGTATCGACCAACTTCACTCCGTTATCATCTAAAGCCACAGAGAGATGGGATTTG 2902
Qy 2962 CTGAGCTCTACATGTTCAAGTGGAGCGCTTCTCATCTCCCTCCGACGACCAATTTGATCA 3021
Db 2903 CAGAACTCTACATCTTCAATGGACAGCTCTTCTCATTTCCACCAACCACTCTACTTGT 2962
Qy 3022 TTAACTGTTTGGTGTCTGCTGGACCTCTTACGCCATCAACAGTGGTGTACCAATCAT 3081
Db 2963 TGAACCTCATAGGCAATTTGGCTGTCTCTTATGCTGTAAACAGTGGCTACCAAGTCTG 3022
Qy 3082 GGGGCGCGCTCTTTTGGGAGGCTCTCTTGTGCTCTCTGCTGATTTGTTCACTTATACCAT 3141
Db 3023 GGGGTCGCGCTTTTTCGGGAGGCTCTTCTTCTGCGCTTATGCGGCTTATGCGCATCTCTACCTT 3082

QY 3142 TCCTCAAGGCTCTATGGCAGCGAACCACACACCGAGGATGTGTCATCGTCTGGGCTG 3201
DB 3083 TCTTGAAGGCTGTGTGGGAAGACAAACCGAACCAATCGTCATTTGCTGTGCTG 3142
QY 3202 TCCTCTCGCTCTATCT 3250
DB 3143 TTCCTCTCGCT 3191

RESULT 3
US-09-221-013A-11
; Sequence 11, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3673
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (71)...(3313)
US-09-221-013A-11

Query Match 37.0%; Score 1341.8; DB 4; Length 3673;
Best Local Similarity 66.2%; Pred. No. 0;
Matches 2097; Conservative 0; Mismatches 992; Indels 80; Gaps 8;

QY 92 GAGGACGGGCGGGGAGCTGTCCAGATCTCGCGCGAGCGGCTGGGCACACAGTTGGA 151
DB 163 GAAGATATGATGGCCAGATATGTGATCTGTGTGATGATGTGGATCTGCTGAAAC 222
QY 152 CGGCGACGCTTCCACCGCCTGCGACGCTGTGCGGCTTCCCGGCTCGCGGCTGCTACGA 211
DB 223 TGGAGATGCTTGTGCGGTGTAATGAATGTGCTTCCCTGTGTGCGGCTTGTCTATGA 282
QY 212 GCACGAGCGCAAGGAGGCAACCGAGGCTGCTCCAGTGTGCAAGCAACCAAGTACAGCGCCA 271
DB 283 GTACGAGAGGAAGATGGAACCTCAGTGTGCTGCTTCCCTCAATGCAAGCTAGATTCAGACGACA 342
QY 272 CAGAGGAGCGCCAGCGATCCGCGGAGGAGGAGCGACACATGATCCGATGATGATG 331
DB 343 CAGGGGAGTCTCTGCTGTGGAAGGATGAAGATGAGATGATGATGATGATGATGATG 402
QY 332 TGACTTCAATACCTCTGATCTGCGCATGAGGACCAAGACGAGAAAGATTCTGACAGGAT 391
DB 403 TGAGTTCAATACGC-----CCAGGAGCTAACCAAGGC 435
QY 392 GCGAGCTGGCGGATGAACACCGGGGGCAGTGGCAATGTTGGCCACCCCAAGTATGACAG 451
DB 436 GAGACACCAACGCGCATGGCGAAGAGTCTTCTTCTCTCTAGACATGAATCTCAACCAAT 495
QY 452 TGGCGAGATCGGCTCTCCAGTATGACAGTGGAGATCCCTAGGGGATACGTCCTTC 511
DB 496 TCCTCTCTCACCCTAGGCGCATACGGTTTCTGAGAGATTCGACGCGCTGATACCAATC 555
QY 512 AGTCACCAACAG-CCAGATGTCAGGAGAAATCCCTGGAGCTTCGCTGTGATCATCATGA 570
DB 556 TGTGCGNACTACATCAGGTCCTTGTGGTCTCTGTGACGAAATGCTATTCTCATCTCCATA 615

QY 571 TGTCCCTACGGGACATCAGCAGCGTGTCCGTTTCCCTATGTGAATCATTCACCAA 630
DB 616 TATTGATCACCAGCAACCTGTCCCTGTAAGAAATCGTGGACCCGTCAAAAGACTT----- 669
QY 631 ATCCGCTCAAGGAGTTCCTCGGCAGTATTGGGAATGTTCCCTCGGAAAGAGAGATTGATG 690
DB 670 -----GAACCTCTTATGGGCTTGTGTAATGTGACTGGAAGAAAGAGATTGAAG 716
QY 691 GTTGAAGAAATGAAGCAGGACAAGGGTGCATTTCCCATGACTAATGGGAACAAGATTGCTC 750
DB 717 GCTGGAAGCTGAAGCAGGAGAGAAATATCTTACAGATGAC-----TGGTA 761
QY 751 CCTCTGAAGTGGCGAGCTACTGACATCGATCTACTGATGATACATACATGGAAGAGC 810
DB 762 AATACCATGAAGGAAAGGAGGAGAAAT---TGAAGGAGACTGGTCCCAATGGGAAGAAC 818
QY 811 CTTTACTGAATGATAAACTCGCCAGCCCTCTATCTAGAAAAGTCCCCATTTCTCTCCA 870
DB 819 TCCAAATGGCTGATGATACACGCTCTTCTATGAGTCGTGTGTGCTATCCCATCTTCTC 878
QY 871 AATAAATCCCTACAGAAATGGTCAATTTCTCGGTTGGTTGTTCTAAGCAATCTTCTCTGC 930
DB 879 GCCTAACCCCTTATCGGGTGTGATTTCTCCGGCTTATCATCTTGTGTTCTTCTCTTGC 938
QY 931 ACTACGCTCTCAAAATCCTGTGCGTAAATGATACCCACTGTGGCTTTTATCTGTTATAT 990
DB 939 AATATCGTCAAACTCACCCCTGTGAAAATGATATCTCTTTGTGGTTGACCTCGGTTATCT 998
QY 991 GTGAGATTTGGTTTGTCTTATCTCTGATACTGGATCAGTCTCCGAGAGTGGTTTCCAATCA 1050
DB 999 GTGAGATCTGGTTTGCATTTTCTTGGCTTCTTGATCAGTTTCCCAATGGTACCCCATTA 1058
QY 1051 ACCGGAGACCTACCTTGTATAGACTGGCTTTTAAAGTATGACCGAGAAAGTGAACCGTCTC 1110
DB 1059 ACAGGAGACTTATCTTGACCGTCTCGCTATAAGATATGATGAGACGGTGAACCATCAC 1118
QY 1111 AGTTGCTGCTGTTGACATATTTGTGACGATGATGACGCTTCCGAGAGTGGAGGAGCCATGCG 1170
DB 1119 AGCTGCTTCTGTTGATGTTGTTGTAGTACAGTGGACCCATTTGAAGAGGCTTCCCTCTG 1178
QY 1171 TCACTCCCAACACTGTGCTATCCATTTCTGCTGCTGATATCCCTGGGACAAGGCTCTCTT 1230
DB 1179 TTACAGCAACACAGTTCTCTCGATTTCTTGTGACTACCGGTAGATATAAGTAGCCT 1238
QY 1231 GCTATGATCTGATGACGAGCTTCAATGCTGACTTTTGAAGCAATTTGGCTGAGACTTCAG 1290
DB 1239 GTTATGTTTCAAGATGATGGTTGAGCTATGCTTACCTTTGAATCCCTTTCTGAAAACGCTG 1298
QY 1291 AGTTTCTAGGAAATGGGTACCACTTTGTGAAGAGTATGACATTTGAACCCAGAGCTCCCG 1350
DB 1299 AGTTTCAAGAAATGGGTACCACTTTGTGAAGAGTATGACATTTGAACCTTAGGGCCCTG 1358
QY 1351 AGTTTACTTTTCCCAAGAAATTTGATTACTTGAAGACAAAGTCCAGCCTTTCATTTGTTA 1410
DB 1359 AATTCTATTTGCCCAAGATAGATTTACTTGAAGCAAGATCCAAACCTCTTTGTTA 1418
QY 1411 AGACCGCGGCGCATGAAGAGAGATATGAAGAAATTTAAATCAGGATTAATGCCCTAG 1470
DB 1419 AAGAGCGAGAGCTATGAAGAGAGATGATGAAGAGTTTAAAGTGAGGATAAATGCTCTTG 1478
QY 1471 TTTCTAAGCATTTGAAGTCCCGGAGAGGATGGATCATGCAAGATGCACACCATGGC 1530
DB 1479 TTGCCAAAGCAAGAAATCCCTGGAAGAGGCTGGACAATGCAAGATGGTACTCCCTGGC 1538
QY 1531 CAGGAACAATACACGGGATCATCTCGGATGATTCAGGTTTTTCTTGTGTCAAGTGGTG 1590
DB 1539 CTGGTAACAACACTAGAGATCATCTTGGATGATACAGGTGTTCTTAGGCCATAGTGGG 1598
QY 1591 GCCTTGATCTAGGGTAAATGAGCTCCCGGTTTATGTTTATGTTGTTCTGTTGAAGCGTC 1650
DB 1599 GTCTGATACCGATGGAATGAGCTGCTAGACTCATCTATGTTTCTCGTGAAGAGCGGC 1658

```
QY 1651 CTGGTCTCCAGCACCAAGAGGCTGGTGCCATGAATGCCCTGTGTTCTGCTCAGCTG 1710
Db 1659 CTGATTTCAACACCAACAAAGAGCTGGAGCTATGAATGCATTGATCGTGTATCTGTTG 1718
QY 1711 TCCTTACTAATGGCAATACATAGTTGTAATCTTGTATTTGATGATCACTACATCAACAAACAGCA 1770
Db 1719 TTCTTACCAATGGAGCATATCTTTTGAACGTGGATTTGTGATCACTTACTTTAATAACAGTA 1778
QY 1771 AGGCTGTCCGAGAAGCTATGCTTCTTAATGGATCCAAACCTAGGTCGCCAAGTCTGTT 1830
Db 1779 AGGCTATTAAGAAGCTATGTTTTCATGATGGACCCGGCTATTGGAAGAAGTGTGCTGCT 1838
QY 1831 ATGTGCAATCCCAAAAGGTTTGATGGGATTTGATAGGAATGATGATGATCAATCAACACAGGA 1890
Db 1839 ATGTCCAGTTCCCTCAACGTTTTGACGGTATTGATTTTGACAGATGATATCCCAACAGGA 1898
QY 1891 ACAGTGTCTTTTGTATTAATTAACCTTGAGGGCCCTTGACGGCATTTCAAGGACAGTTTATG 1950
Db 1899 ATATAGTCTTTTTCGATATTAACATGAAGGGGTTGGATGGTATCCAGGGTCCAGTATATG 1958
QY 1951 TGGAACTGGTGTGTTTCAACAGAACAGCTATCTATGTTATGAGCC-----CCCA 2003
Db 1959 TGGTACTGTTGTTTATAGGAGGCTCTATATGGGTATGATGCTGTTTGTAGCG 2018
QY 2004 ATTAAGCGAAGAACCCAGGTTTCTTGCAATCA--CTATGTGGGGCGAAGAAAGGCA 2061
Db 2019 AAGAAGATTTAGAACCATAATATTATGTCAAGAGCTGTGGCGGTCAAGGAAGAAAGTA 2078
QY 2062 GCAAGTCAAGAAAAGGAGCTCAGATGAAGAAAAGTCAACAAAGCATGTGACAGATTTCTG 2121
Db 2079 AAGTAGCAAGAATATAACTTACCAAAAGAGGAGGAGCATCAACAGAAAGTCACTCCAATG 2138
QY 2122 TTCCAGTATTCAATCTCGAAGACATAGAGGAGGTTGTAAGGTGCTGGGTTTCATGATG 2181
Db 2139 CTCACATTTCAATATGAGAGCATCGATGAGGGTTTGAAGT-----TATGATGATG 2192
QY 2182 AGAAATCAGTTCTCATGCTCAAAATGAGCTTTAGAGAAGAGATTTGGCCAGTCAGACGAT 2241
Db 2193 AGAGTCTATTCTAATCTCCAGAGGAGTGTAGAGAAGCGTTTGGTCAGTCGCGGTAT 2252
QY 2242 TTGTTGCTCCACTCTGATGAATATGTTGTTCTTCAGTCTCCACTCCAGAAATCTC 2301
Db 2253 TTAATGCGGCAACCTTCATGGAACAGAGCGGCAATTCACCAACAAACAAATCCCGCTACTC 2312
QY 2302 TTTTGAAGAAGCTATCCATGTCTATAAGTTGTGCTATGAGCAAGTCTGAATGGGAA 2361
Db 2313 TTCTGAAGAGGCTATTCTATGTTATAGTGTGTTACGAGACAAAGTGAATGGGCA 2372
QY 2362 CTGAGATTGGTATGATCTATGATCTGTCAAGAGATATTCTTACTGGATTCAGATGC 2421
Db 2373 AAGAGATTGGTATGATCTATGTTCCGTGACGGAAGATATTCTTACTGGGTTCAAGATGC 2432
QY 2422 ACGCAAGAGCTGCGTTTCACTCTATTGTCATGCCCAAGCGCCAGCTTTCNAGGGATCTG 2481
Db 2433 ATGCCCGGGTTGATATCGATCTACTGCAATCTCCACGCGCTGCGTTCAAGGGATCTG 2492
QY 2482 CCCCCTCAATCTTTTCAGATCGTCTGAACCAAGTCTGCGGTGGGCTCTCGTTTCTGTTG 2541
Db 2493 CACCAATCAATCTTTCTGATCGTTTGAACCAAGTCTTCGATGGGCTTTGGATCTATCG 2552
QY 2542 AAATCTTTTTCAGCGGATTTGCCCTTATGATNGCTAGCGAGGCGGCTCAAGTTCC 2601
Db 2553 AGATTCCTCTTAGCAGCATTTGCTATCTGTTATGTTTACCTATGGAAGGTTGAGACTTT 2612
QY 2602 TGGAGATTTGGTTTACATCAACACCACTTTTACCCACTAACTCTCTCCCGTTCTAG 2661
Db 2613 TGGAGAGATTCGCTTTATCAACACCATCTGTTATCTCTATTTACATCATCCCTCTTATG 2672
QY 2662 TCTATTGTATTTGCCGTCTATCTGCTGTCTCACTGGAAAAGTTTCATCATGCCAGAGATTA 2721
Db 2673 CGTATTGTATTTCTCCCGCTTTTGTCTCATACCGACAGATTTCATCATACCCGAGATA 2732
QY 2722 GCAACTGGCCAGTATCTGGTTTATGCGGCTCTTCTCTTCAATTTTCGCCACTGGTATCC 2781
```

```
Db 2733 GCACTAGCGAGTATTGGTTTCTTACTCTCTCACTCAATTCGCTGTGACTGAAATCC 2792
QY 2782 TTGAGATGAGGTGAGTGGTGTGGCATTTGACGATGGTGGAGAAAGAAAGTTCTCTGG 2841
Db 2793 TGGAGCTGAGATGGAGCGGTGTGAGCATTTGAGGATTTGGTGAGGAAACGAGCAGTTCTGG 2852
QY 2842 TCATTGGAGGTATCTCTGCAATCTGTTTTCGCTCTTTCAGGGTCTTCTGAAGGTGCTTG 2901
Db 2853 TCATTGGGCAATCCGCCCATCTTTTGTGCTCTTCCAAGGTCTACTTAAGGTCTCTTG 2912
QY 2902 CCGGTATCGACACCAACTTCACTGTCACTCAACCTCAAGGCTTAATGACGAAGAAGCGCACTTTG 2961
Db 2913 CTGATATCGACACCAACTTCACTGTCACTCAACCTCAAGGCTTAATGACGAAGAAGCGGATTTG 2972
QY 2962 CTGAGCTCTACATGTTCAAGTGGAGAGGCTTCTCATCCCTCCAGACGACCAATTTTGATCA 3021
Db 2973 CAGAACTCTACATCTTCAAAATGGAGCTCTTCTCAATTCACCAACCAACCGTCTCTACTTG 3032
QY 3022 TTAACATGGTGTGTTGCTGGCACCTCTACGCGCATCAACAGTGGTTTACCAATCAT 3081
Db 3033 TGAACCTCATAGGATTTGTGGCTGTCTCTTATGCTGTAAACAGTGGCTACCAAGTCTG 3092
QY 3082 GGGGGCCGCTCTTTTGGAAAGCTCTTCTTTCGCTTCTGGGTGATTTGTTCACTTATACCCAT 3141
Db 3093 GGGGTCCGCTTTTCGGGAAGCTCTTCTTTCGCTTATGGGTTATTTGCCCATCTCTACCCCT 3152
QY 3142 TCCTCAAGGCTCTTATGGGAGGCAAAACCGCACACCGAGATTTGATCGTCTGGGCTG 3201
Db 3153 TCTTGAAGGCTCTTGGGAAGCAAAACCGAACCAACCAACCATCGTCAATTTGTTGGTCTG 3212
QY 3202 TCCTCTCGCTCTTATCTCTCTTCTGCTGGGTTGCTGTTGATTCCTT 3250
Db 3213 TTCTTCTCGCTCCATCTTCTGCTTGGTTCAGATCAATCCCTT 3261

RESULT 4
US-08-960-048-1
; Sequence 1, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; CURRENT FILING DATE: 1997-10-29
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3328
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-08-960-048-1

Query Match 33.5%; Score 1215.6; DB 3; Length 3328;
Best Local Similarity 59.2%; Pred. No. 0;
Matches 1710; Conservative 0; Mismatches 739; Indels 21; Gaps 3;

QY 807 GAGCCTTTACTGAATGATGAACCTCGCCAGCCTCTATCTAGAAAGTCCCAATGCTTCC 866
Db 617 GAAGATAAACCGGACCGGATGCTTCCAGCCCTCTCGACTATAATTCCAATCCCGAAA 676
QY 867 TCCAAATAAATCCCTACAGAATGGTCAATTTGTTCTGCGGTGGTGTGTTCTAAGCATCTTC 926
Db 677 AGCAGACTTGCAACCATACCGAACCGGATCATTTATGCGATTGATCATCTTTGGTCTTTTC 736
QY 927 CTGCACTACGCTCTCAGAAATCTGTGGCGTAAATGATACCCACTGTGGCTTTTATCTGTT 986
```

Db 737 TTCCATTATCGAGTAACAAACCCGGTTGACAGTGTCTTTTGGACTGTGGCTCACTTCAGTC 796
QY 987 ATATGTGAGATTGGTTTTCCTTTATCTCGGACTGATGATCAGTTCCTCCGAAAGTGGTTTCCA 1046
Db 797 ATATGTGAAATCTGGTTTGCATTTTCTGGGTGTTGGATCAGTTCCTTAAGTGGTATCCT 856
QY 1047 ATCAACCGGGAGACCTACCTTCATAGACTGGCTTTAAAGGTATGACCGAAGGTGAACCG 1106
Db 857 GTTAACAGGGAACAATACATTCACAGACTATCTGCAAGATATGAAGAGAGGTGAACCT 916
QY 1107 TCTCAGTTGGCTGTCTTGGACATATTTGTCAGTACAGTCGACCCCTTGAAGAGGCCACT 1166
Db 917 GATGAACCTTGTGCGAGTTGACTTCTTCTGAGTACAGTGGATCCCATTTGAAGAGGCTCCA 976
QY 1167 ATCGTCACTGCCAACACTGTGTATCCATTCTTGTGTTGATATCCCGTGGACAGGTC 1226
Db 977 TTGATTACTGCCAAATCTGTGTCTTCCATTCCTTGGCTTGGACTACCCGGTGGATAGGTC 1036
QY 1227 TCTTGTATGTATCTGATGACGGAGCTTCAATGCTGACTTTTGGACATTTGCTGAGACT 1286
Db 1037 TCTTGTATATATCTGATGATGGTGGGCCATGCTGACATTTGAATCTCTAGTAGAACA 1096
QY 1287 TCAGAGTTTGTAGGAAATGGGTACCATTTGTGAAGAGTATGACATTTGAACCCAGAGCT 1346
Db 1097 GCCGACTTTGCAAGAAAGTGGGTTCATTTCTGCAAAAAATTTTCCATTTGAACCCCGGCA 1156
QY 1347 CCCGAGTTTACTTTTCCAGAAATTTGATTTACCTGAAAGACAAAGTCCAGCTTCATTT 1406
Db 1157 CTTGAGTTTACTTCTCACAGAAGATTGATTAATCTTGAAGATGAAGTGCAGGCCCTCTTT 1216
QY 1407 GTTAAAGACCGCGGCCATGAAGAGAGATATGAAGATTTTAAATCAGGATAAATGCC 1466
Db 1217 GTAAAGAACGTAGAGCTATGAAGAGATTTAAGAGTACAAATTCGATCATGCT 1276
QY 1467 CTAGTTTCTAAGGCATTTAAAGTCCCGAGGAGGATGGATCATGCAAGATGGCACACA 1526
Db 1277 TTAGTTGCAAGGCTCAGAAAAACCTGATGAAGATGCACAAATGCAAGATGCAACTTCT 1336
QY 1527 TGGCAGGAAACAATACAGGAGTATCTCTGAAATGATTCAGGTTTCTTGGTCAAGT 1586
Db 1337 TGGCCAGGAAATAACCGCGGTGATCACCCCTGGCATGATTCAGGTTTCTTGGATATAGT 1396
QY 1587 GGTGGCTGTGATGAGGATTAAGTCCCGGTTTATGTTTATGTTCTCGTGAAGAG 1646
Db 1397 GGTGCTGTGACATCGAAGAAATGAATCTCTCGACTGGTTTACGTCTTAGAGAGAAG 1456
QY 1647 CGTCTGGGTTCCAGCACCAAGAGGCTGGCCATGAATGCCCTTGTTCGTGTCCTCA 1706
Db 1457 AGACCTGGCTACCAACACCAAAAAAGGCTGGTGTGCTGAAAAATGCTTTGGTTAGGTTGTCT 1516
QY 1707 GCTGTCTTACTAATGGAACAATACATGTTGAATCTTGATTTGATGATCACTACATCAACAAC 1766
Db 1517 GCAGTCTTTACAAATGCTCCCTTCATCCTCAATCTTGATGTGACCACTATGTTAAACAAT 1576
QY 1767 AGCAAGCTGTCCGAGAAGCTATGCTCTTCTTAATGATCCAAACCTAGCTCCGCAAGTC 1826
Db 1577 AGCAAGCAGTTAGGGAGGCAATGTCTTCTTGATGGAACCAAGAGTTGGTCAGATGTA 1636
QY 1827 TGTATTGTGCAAGTCCCAAAAGGTTTGTGAGGATTTGATAGGAATGATCGATATGCAAAAC 1886
Db 1637 TGCTATGTGCAAGTTTCTCAAGATTTGATGGCATAGATAGGAGTATCGATATGCCAAT 1696
QY 1887 AGGAACAATCTTTTGGATATTAATCTTAGAGGGGCTTGAACGGCAATTCAGAGCAAGTT 1946
Db 1697 AGGAACACAGTTTCTTGTATGTTAATCATGAAGGTCCTTGATGGAATCCAAAGGGCCAGTT 1756
QY 1947 TATGTGGGAACGTGGTTGTTTTCACAGAACACTATCTATGTTTATGAGCCCCCAATT 2006
Db 1757 TATGTGGGAACAGTTGTGTTTTCATATAGCAAGCACTTTATGGCTATGTTCCACCTTCA 1816
QY 2007 AAGCGGAAGAGCGAGGTTTCTTGGCATCACTATGTGGGGGCAAGAGAGCAAGCAAG 2066
Db 1817 ATGCCAAGTTTCCCAAGTCATCTCTCATCTTGTCTGTGTTGCTGCCCGGGCAAGAG 1876

QY 2067 TCAAGAAAAGGAGCTCAGATAAGAAAAGTCGAACAGCATGTGCAGAGTTCTGTGTCCA 2126
Db 1877 GAACCTAAAGATCCATCAGAGCTTTATAGGATGCAAAACGGGAAGAACTTATGTCGTCC 1936
QY 2127 GTATTCAATCTCGAAGACATAGAGGAGGTGTGAAGGTGCTGGGTTTGATGATGAGAAA 2186
Db 1937 ATCTTTAACCTTAGGGAATTGACAATTATG-----ATGAGTATGAAGA 1981
QY 2187 TCAGTTCTCATGCTCAAAATGAGCTTAGAAGAGAAATTTGGCCAGTCAGACGACATTTGTT 2246
Db 1982 TCAATGTTGATCTCTCAAAACAGCTTTGAGAAAACCTTTTGGCTTATCTTCAGTCTTCATT 2041
QY 2247 GCCTCCACTCTGATGGAATATGGTGTCTCTCAGTCTCTCAGTCTCCACCTCCAGAACTCTTTTG 2306
Db 2042 GAATCTACACTAATGGAGATGGAGAGTGGCTGAATCTGCCAACCTTCCACACTAATC 2101
QY 2307 AAAGAAGCTATCCATGTCATAAAGTTGTGGCTATGAGGACAAAGTCTGAATGGGGAACAG 2366
Db 2102 AAGGAAGCAATTCATGTCATCAGCTGTGGCTATGAAGAGAGACTGCATGGGGGAAGAG 2161
QY 2367 ATTGTTGGATCTATGGATCTCTCAGAGATATTTCTTACTGGATTAAGATGCAAGCA 2426
Db 2162 ATTGATGGATATATGGTTTCACTCAGGATATCTTAAACGGGCTTCAAAATGCACCTGC 2221
QY 2427 AGAGCTGCGGTTCACTATTTGCAATGCGCCAGCGCCAGCTTTCAGAGGATCTGCCCC 2486
Db 2222 CGAGGATGGAGATCGAATTTACTGCAATGCCCTTAAAGCCAGCATTCAGAGATCTGCACCC 2281
QY 2487 ATCAATCTTTCAAGATCTCTGAAACCAAGTGTGCGGTGGGCTCTCGGTTCTCTGAAAT 2546
Db 2282 ATCAATCTGTCTGATCGGTTGCACCAAGTTCTTCGATGGGCTCTTGGATCTGTGAAAT 2341
QY 2547 CTTTTCAGCGCGCATTTGCCCTTATGGTATGCTACGGA---GGCGCCTCAAGTTCCGTG 2603
Db 2342 TTCTTAAGCAGGCAATTTGCCCTCTATGGTATGGCTTTTGGAGGTGGTCTTAAATGGCTT 2401
QY 2604 GAGAGATTCGCTTACATCAACACCACTTTTACCACCTTAACCTCTCTCCGCTCTTAGTC 2663
Db 2402 CAAAGACTAGCATATATAACACCACTGTCTATCTTTTCCATCTCCCTCCACTCATGGC 2461
QY 2664 TATTGTATATTCCCTGCTATCTGTCTGCTCACTGGAAAGTTTCATATGCGACAGATTTAGC 2723
Db 2462 TATTGTTCACTACAGCAATCTGTCTCTCAAGGAAATTTATCATACCAAGCTCTCA 2521
QY 2724 AACCTGGCAGTATCTGGTTTCAATTTTCCCTTTTCAATTTTCCGCACTGGTATCCTT 2783
Db 2522 AACCTGGCAAGTGTCTCTTCTTGGCCTTTTCTTCCATTTATCGTACTGTGTCTC 2581
QY 2784 GAGATCAGGTGAGTGGTGTGGCATTTGACAGTGTGTGAGGAAGTGAACAGTTCTGGGTC 2843
Db 2582 GAGCTCCGATGAGTGGTGTGAGCACTTGAAGACTTATGGCGTAAACAGCAGTTTGGGTC 2641
QY 2844 ATTGGAGGTATCTCTGCACATCTGTTTGGCGCTTTTTCAGGGCTCTTCTGAAGGTGCTTGC 2903
Db 2642 ATCGTGGGTTTTCAGCCATCTCTTGGCGCTTCCAGAGTTTCTTAAAGTGTGCG 2701
QY 2904 GGTATCGACCAACTTCTCATCTGTCACTCAAGGCTTAATGACGAAGAGGCACTTTGCT 2963
Db 2702 GGCATTGACACCAACTTTACTGTCTGTCGCAAGCAGCTGATGATGCA---GATTTTGGT 2758
QY 2964 GAGCTCTACATGTTCAAGTGGAGCGGCTTCTCATCTCCCTCCGACGACCAATTTTGATCAT 3023
Db 2759 GAGCTCTACATGTTGAAATGGACTACATTTCTAAATCCCTCCAAACAACTCTCTCATCTC 2818
QY 3024 AACATGTTGGTGTCTGCTGGCACTCTTACGCCATCAACAGAGTGGTTTACCAATCATGG 3083
Db 2819 AACATGTTGGTGTCTGTCGCGGATTTCTCCGATGCCCTCAACAAAGGTACGAGCTGG 2878
QY 3084 GGGCGCTCTTTGGGAAGCTCTTCTTGGCTTCTGGGTGAATTTGCTTATACCATTC 3143
Db 2879 GGACCACTCTTTGGCAAGTGTCTTCTTCTGCGTCTGCGTCTCTCATCTTATTCATTC 2938

QY 2367 ATTGGTTGATCTATGAGTCTGTGACAGAGATATTCTTACTGGAATCAAGATGCACGCA 2426
|||
Db 2162 ATTGNTGATATATGTTTCACTGATCTGAGGATATCTTAACGGCTTCAAAATGCATGC 2221
|||
QY 2427 AGAGGTCGCGTTTCACTTATGATGCCCCAGCCGCTTCAAGGATCTGCCGCC 2486
|||
Db 2222 CGAGGATCGAGATCGATTTACTGATGCCCCCTAAAGCCAGCATTCAGAGATCTGCACCC 2281
|||
QY 2487 ATCAATCTTTCAGATCGTGTGACCAAGTGTGCGGTGGCTCTCGTTCTGTGAAAT 2546
|||
Db 2282 ATCAATCTGTCTGATCGGTGTGACCAAGTGTCTCGATGGCTCTGTGAAAT 2341
|||
QY 2547 CTTTTAGCGCGCATGCCCCCTATGATGATGCTACGGA---GGGCGCCTCAAGTTCCCTG 2603
|||
Db 2342 TTCTTAAGCAGGATTTGCCCTCTATGATGCTTTGGAGTGTCTTAAATGGCTT 2401
|||
QY 2604 GAGAGATCGCTTACATCAACACCACTATTAACCACTAACCCTCTCTCCGCTTTAGTC 2663
|||
Db 2402 CAAAGACTAGCATATATAAACACCACTTGTCTATCTCTTACATCCCTTCCACTCATTCG 2461
|||
QY 2664 TATTGATATGCTGCTATCTGTCTGCTCACTGGAAGTTTATCATGCGCAGATTAGC 2723
|||
Db 2462 TATTGTTACTACCAAGCAATGTCTCTCTCACAGGAAATTTATCATCAACGCTCTCA 2521
|||
QY 2724 AACTGGCCAGTATCTGGTTTCAATGCGCTCTCTCTTCAATTTTCCCACTGGTATCCTT 2783
|||
Db 2522 AACCTGGGAGTGTCTCTTTTGGCCCTTTCTTCCATTAATGCTGCTGTCTC 2581
|||
QY 2784 GAGATGAGTGTGAGTGTGTGGCATGAGTGTGAGGAGTGAACAGTTTCTGGTC 2843
|||
Db 2582 GAGCTCCGATGAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2641
|||
QY 2844 ATTGAGGATCTCTGCACTCTGTTTGGCTTTTCCAGGCTTCTGAGGCTTCTGCC 2903
|||
Db 2642 ATCGTGGCGTTTTCAGCCCATCTCTTTGGCGCTTCCAGGTTTCTTAAAGTGTCTGG 2701
|||
QY 2904 GGTATCGACACCAACTTCACTGTCACTCAAGGCTTAATGACGAGGAGGCGACTTTGCT 2963
|||
Db 2702 GGCATTTGACACCACTTTTACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2758
|||
QY 2964 GAGCTCTACATGTTCAAGTGTGAGGAGTGTCTTCTATCCCTCCGACGACCAATTTGAT 3023
|||
Db 2759 GAGCTCTACATGTTGAAATGGAATGATCTTCTTAAATCCCTCCAAACACTCTCTCATG 2818
|||
QY 3024 ACATGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3083
|||
Db 2819 ACATGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2878
|||
QY 3084 GGGCGCTCTTTGGGAAGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3143
|||
Db 2879 GGACCACTTTTGGCAAGTGTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2938
|||
QY 3144 CTCAGGGTCTTATGGGAGGCAAAACCGCACCGGATGCTATGCTGCTGCTGCTGCTGCTG 3203
|||
Db 2939 CTCAGGGTCTTATGGGAGGCAAAACCGCACCGGATGCTATGCTGCTGCTGCTGCTGCTG 2998
|||
QY 3204 CTCCTGCTTCTATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3263
|||
Db 2999 TTGTTGGCTTCTGCTTCTCTCTTCTGTTGGTTCCGATCAACCCGTTTGTGACACCGCC 3058
|||
QY 3264 GCTGGGCCAA 3273
|||
Db 3059 GATAGCACCA 3068
|||

RESULT 6

US-09-221-013A-7
; Sequence 7, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.

; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 3828
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (239)..(3490)
US-09-221-013A-7

Query Match 32.4%; Score 1176.6; DB 4; Length 3828;
Best Local Similarity 68.2%; Pred. No. 2.2e-311;
Matches 1667; Conservative 0; Mismatches 764; Indels 12; Gaps 2;
QY 817 TGAATGATGAAACTGCCAGCCTCTATCTAGAAAAGTCCCATTTGCTTCTCCCAATAA 876
|||
Db 999 TGATGATGAAAGAGAACAACTCTCTCAAGAAAGTACCTATTCGTTCAAGCAGATTA 1058
|||
QY 877 ATCCCTACAGAATGGTCAATTGTTCTGCGGTTGGTTGTTCTTAAGCACTTCTCTGCACTAC 936
|||
Db 1059 ATCCCTACAGGATGTTAATTCGTGTCGCTCGCATCTCTGCTCTTCTTCAATATA 1118
|||
QY 937 GTCCTCAAAATCCTGTGCGTAAATGATACCCACTGTGGCTTTTATCTGTATATGAGA 996
|||
Db 1119 GAATTCCTCCATCCAGTCAATGATGATATGATGATGATGATGATGATGATGATGATG 1178
|||
QY 997 TTTGGTTTCTTTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1056
|||
Db 1179 TATGTTTTCAGTGTCTTGGATCTTGTATCAATTTCCCAATGATGATGATGATGATG 1238
|||
QY 1057 AGACCTTACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1116
|||
Db 1239 AAACATACCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1298
|||
QY 1117 CTGCTGTGATATATTTGTCAGTACAGTCCGCTTGAAGAGCCACTATGCTGCTGCTG 1176
|||
Db 1299 CACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1358
|||
QY 1177 CCAACACTGCTGATCTTCTTCTGATGATGATGATGATGATGATGATGATGATGATG 1236
|||
Db 1359 CAAACACAGTCTTTCCTCATCTAGCATGATGATGATGATGATGATGATGATGATGATG 1418
|||
QY 1237 TATCTGATGACGAGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1296
|||
Db 1419 TATCAAAACAATGTCAGCTATGCTTACATTTTGAAGTCTCTCTGATACAGCTGATTTG 1478
|||
QY 1297 CTAGGAAATGGGTACCATTTGTCAGAGATATGATGATGATGATGATGATGATGATG 1356
|||
Db 1479 CTACAAATGGGTTCCTTTTCTGTAAGATTTAATATCGAGCCACGAGCTCTGCTGCTG 1538
|||
QY 1357 ACTTTTGGCAGAAAATGATTAACCTGAAAGACAAAGTCCAGCTTCAATTTTAAAGACC 1416
|||
Db 1539 ATTTTCTCAGAAGATGATTAACCTGAAAGACAAAGTTCATCTGCTTTTGTGACGGAAC 1598
|||
QY 1417 GCCGGGCCATGAAGAGAGATATGAAGATTTTAAATCAGGATAAATGCCCTAGTTTCTTA 1476
|||
Db 1599 GTCGTGCTATGAAGAGATATGAAGATTTTAAAGTGAAGATAAATGCTGTTGCTA 1658
|||
QY 1477 AGCATTTGAAAGTCCCGAGGAGGATGATGATGATGATGATGATGATGATGATGATG 1536
|||
Db 1659 CTGCACAGAAAGTGGCTGAGGAACTGATGATGATGATGATGATGATGATGATGATG 1718
|||
QY 1537 ACAATACCAGGATCATCTCTGGAATGATTCAGGTTTCTCTGCTCACAGTGGCTGCTG 1596
|||

Db 1719 ACAACGTCCTGGACATCTCGAATGATTCAGGTGTTCTTGGGTCTATAGTGGAGTTCGTG 1778
QY ATACTGAGGGTAATAGAGTCCCGCTTTAGTTTATGTTCTCGTGAAGAGTCTCGGGT 1656
Db 1779 ATACGGATGTAATAGAGTTACACAGTCTAGTGTATGTTCTCGTGAAGAGCGGCTGGAT 1838
QY TCCAGCACCAAGAAGCGTGGTCCCATGAATGATGCTTGTGTTCTCGTGTCTCAGCTGTCCTTA 1716
Db 1839 TTGATCACCAAGAAGAGTGGAGCTATGAATTCCTTGATCCGAGTCTCTGCTGTTCTAT 1898
QY CTAATGACCAATATAGTGTGAATCTTGATTTGATCACTACATCAACAAACAGCAAGGCTG 1776
Db 1899 CAAACGCTCTTACCTTCTTAATCTGATTCGATCACTACATCAACAAACAGCAAGCA 1958
QY TCCGAGAGCTATGCTGCTTCTAATGATCAAAACCTAGTTCGGAAGTCTGTTATGTC 1836
Db 1959 TTAGAGAACTATGTTGTTTATGATGACCCGCAATCGGGAAGAAAGTTTGTATGTC 2018
QY AGTTTCCCAAAAGTTTATGAGGATTCATAGGAATGATCGATATGCAAAACAGGAACACTG 1896
Db 2019 AGTTTCCGACAGATTTATGAGGATTTATAGACATGATAGTACTCAACCGTAACTG 2078
QY TCTTTTGTATTAATCTGAGGGCCCTTGAGGCACTTCAAGGACCACTTTATGTGGAA 1956
Db 2079 TGTCTTTGATATTAACATGAAGGCTCTTGATGGATACAGGACCGATATATGTCGGGA 2138
QY CTGTTGTGTTTCAACAGAAACAGCTATCTATGTTATGATGACCCCGCAATTAAGCGAAGA 2016
Db 2139 CAGTTGTGTTTATGAGAAACAGGCTCTTTATGTTTGTATGACCAAAAGAAAGAAAC 2198
QY AGCCAG-----GTTTCTTGCACTACATCTGTTGGGGCAAGAAAGGCAAGCAAGT 2067
Db 2199 CACAGGCAAAACCTGTAACTGTTGGCTTAAAGTGTGTTGTTGTTGGTTGAA 2258
QY CAAGAAGAGGAGCTCAGATAAGAAAAAGTCAACAGCAATGTGGACAGTTCTGTTCCAG 2127
Db 2259 AGAAGAGTAAACGAAAGCCACAGATAAGAAACATAACATAAGAGACTTCAAAAGCAGA 2318
QY TATTCATCTCGAAGCATAGAGAGGTTGTTGAAGTGTCTGGTGTGATGATGAGAAT 2187
Db 2319 TTATGCGCTAGAGAATGTGCAGGAAGTGTATCTGCTCCAGTGTCAAAATGTTGAGAAGA 2378
QY CAGTTCTCATCTCTCAATAGCTTAGAGAAGATTTGGCCAGTCAAGCAGCATTTGTTG 2247
Db 2379 GATCTGAAGCAACATTAATGAATTTGAGAGAAGTTTGGACATCTCCGTTTTCGTTG 2438
QY CTTCCACTCTGATGGAATATGTTGTTCTCTCAGTCTCTCACTCCAGAAATCTCTTTGA 2307
Db 2439 CTTCTGCTTCTACAGAACGTTGGAGTTCCCGCTAACGCAAGCCCGCATGTTTGTAA 2498
QY AAGAGCTATCCATGTATAGTTTGGCTATGAGGCAAGTCTGAATGGGAACTGAGA 2367
Db 2499 GAGAAGCAATTCAGATTATAGTGGCGGTACCAAGATAAAACCGAATGGGGAAGAGA 2558
QY TTGTTGGATCTATGATCTCTACAGAAGATATTTCTACTGATTCAGATTCAGATGACGCA 2427
Db 2559 TCGGTGATTTATGATCGTGTACTGAGATATCTGACGGTTTCAAGATGATTCGCC 2618
QY GAGGCTGCGTGTATGATTCATGATGCAAGCGCCAGCTTTCAAGGATCTGCGCCCA 2487
Db 2619 ATGGATGAGATCTGTGTACTGTATGCTTAAGCGTGCAGCTTTTAAAGGATCTGTCCTA 2678
QY TCAATCTTTCAGATCTGTGAACCAAGTCTGCGGTGCTCTGTTCTGTTCAAAATTC 2547
Db 2679 TTAATCTTCAGATCTGTTTCAATCAAGTTTACGTTGGGCTCTTGGCTCTGATGAGATTT 2738
QY TTTTCAGCGGATTCGCCCTTTATGTTATGCTACGAGGCGGCTCAAGTTCTCTGGAGA 2607
Db 2739 TCTTCAGCAGACATTTGTCGATATGGTATGTTATGTTGTTGTTTAAATGTTGGAGA 2798
QY GATTGCGTTATCAACAAACCAACATTTAACCACTAACCTCTCTCCCGCTTCTAGTCTATT 2667

Db 2799 GATTCTCTTACATCAACTCTGTCGTCTATCTCTTGGACTTCACITTCCTCATCTGCTATT 2858
QY GTATATTCCTGCTATCTGCTGCTCACTGGAAGTTTCATCATGCCAGAGATTAGCAACT 2727
Db 2859 GTTCTCTCCCGGGTGTGTTTACTACAGAAAAATTCATGCTCCCTGAGATAAGCAACT 2918
QY TGGCCAGTATCTGTTTCAATTCGCTCTCTCTTCAATTTTTCGACACTGGTATCTCTGAGA 2787
Db 2919 ACGAGGTATCTCTTCACTGCTCATGTTTATATCCATAGCAGTAACTGGAATCTCGAAA 2978
QY TGAGTGGAGTGTGTTGGCATTTGACGAGTGGTGGAGAAATGAACAGTCTGGGTCTATTG 2847
Db 2979 TCGAATGGGAGGTGTCGGAATCGATGATTGGTGGAGAAACGAGCAGTTTTTGGGTAAATCG 3038
QY GAGTATCTCTGACATCTGTTGCGCTCTTTCAGGGTCTTCTCAAGGTCTTCTGAGGTCTCCCGGTA 2907
Db 3039 GAGGGGCTCTCTGCACTATTTGCTCTGTTTCAAGGTTTGTCTCAAGTTCTAGCCGAG 3098
QY TCGACACCAACTTCACTGTCACTCAAGGCTATGACGAAGAGGCGACTTTTCTGAGC 2967
Db 3099 TTAACACGAATTTCAAGTCACTTCAAAAGC---AGCAGACGATGGAGCTTCTCTGAGC 3155
QY TCTACATGTTCAAGTGAACGCGCTTCTATCCCTCCGACGACCATTTTGTATCATTAACA 3027
Db 3156 TTTACATCTTCAAGTGAACAACTTTGTTGATTTCTCCGACAACTTCTGTATCAATAACA 3215
QY TGGTGTGTCGTTGCTGGCACCTCTACGCCATCAACAGTGGTTACCAATCATGGGGC 3087
Db 3216 TCAATGAGTATTTGTCGGGCTTCTGATGCCATTAGCAATGGCTATGACTCATGGGAC 3275
QY CGCTCTTTGGGAAGCTCTCTTTTGGCTTCTGGGTGATTGTTCACTTATACCCATTCTCTCA 3147
Db 3276 CTCTTTGGGAGACTTTTCTCTGCTCTTTGGGTCAATTTTCAATTAACCACTTCTCTCA 3335
QY AGGCTCTTATGGGAGCAAAACCGCACCGGAGATGTCATGCTGGGCTGCTCTCC 3207
Db 3336 AGGGAATGCTTGGGAAGCAACAAAATGCTACGATTTATGTTGCTGCTGTTCTTCTTC 3395
QY TCGCTTCTATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3250
Db 3396 TAGCTTCGATCTTGACACTCTTGTGGGTGAGAAATTAACCCGTT 3438

RESULT 7

US-09-221-013A-1
; Sequence 1, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2448
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1887)
US-09-221-013A-1

Query Match 27.1%; Score 984.2; DB 4; Length 2448;
Best Local Similarity 72.1%; Pred. No. 7.1e-259;

; SEQ ID NO 13
; LENGTH: 1741
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1741)
US-09-221-013A-13

Query Match 15.7%; Score 568.8; DB 4; Length 1741;
Best Local Similarity 62.0%; Pred. No. 2.7e-145;
Matches 1031; Conservative 0; Mismatches 547; Indels 84; Gaps 5;

```
QY 53 CCGCGCATGACGCGACGCGGACGCGCTGAGTCCGGGAGGACACGCGGCGCGGAGCGT 112
DB 160 CCGCGCGACGCGGACGCGGACGCGGCTTAAGCGAGGAGAGAGTGTGAATGTTCAGGT 219
QY 113 GTGCCAGATCTGCGCCGACGCGCTGGGACACACGCTTGACGCGGACGCTTTCACCGCTG 172
DB 220 CTGCCAGATTTGTGGGACACTGTGTGGGCTCTCGGCCACCGCGGACGCTTTTGTGCGTG 279
QY 173 CGAGCTGTGCGGCTTCCGGTCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 232
DB 280 CAATGAGTGGCGCTTCCCGCTCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTG 339
QY 233 CCAGGCTGCTCCAGTGCAGACCAAGTACAGCGCCACAGCGGAGGCGGCGGATCCG 292
DB 340 CCAGTGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 399
QY 293 CGGCGAGGAGGCGGACGACACTGATGCGGCTGATGATGATGATGATGATGATGATGATG 352
DB 400 GGGCGATGAGGAGAGAGATGTTGATGACCTGGACATGAATTCATTA----- 450
QY 353 TGGCACTGAGGACGAGAGAGAGAGATGCTGAAGGATGCGGCTGCGGCTGCGGCTGCGG 412
DB 451 -----TAAGCATGG 459
QY 413 CGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATGGCGCTCTCCAA 472
DB 460 CAATGGCAAGAGTCCAGTGGCGAGATACAGAGACAGGGGAGAGATGTTGACCTGTCTTC 519
QY 473 GTATGACAGTGGAGAGATCCCTAGGGGATAGTCCCTTCAGTACCAACAGCGGAGATGC 532
DB 520 ATCTCTCGCCACGAAACATCGAATCCCGCTGACAAAGTGGCA---ACAGATCTC 576
QY 533 AGGAGAAATCCCTGAGGCTTCCGCTGATCATCATGATGTCCTCCCTACGGGAGCATCAG 592
DB 577 AGGAGAGATCCCTGATGCTTCCCGCATCGCCATTTCTATCCGCGAGCGGAAACATCAAGCTA 636
QY 593 CAGAGCTGCTCGGTTTCCCTATGTAATCATTCACCAATCCGTCAAGGGAGTTC---TC 649
DB 637 TGTGATCAAGTGTTCAGTTCCTGTGAGGATTTGGACCCCTCAAGGACTTGAATTC 696
QY 650 CGGCGAGTATGGGAATGTTGCTGGAAAGAGAGAGTGTGATGCTGGAAATGAAGACGGA 709
DB 697 CTATGGGATTAACAGTGTGACTGCGCAAGAAAGAGTTCAGCTGGAGGAAACAGACGGA 756
QY 710 CAAGGTCGATTCCTGATGATGAGTGGACAGCAATGCTCCCTCTGAGGTGCGGGAGC 769
DB 757 CAAAAATATGATGACGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 807
QY 770 TACTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 829
DB 808 CATGA-----AGGACTGGTTCAAATGGTGAAGATATCCAAATGGTGTGATGATG 858
QY 830 TGGCAGGCTCTATCTAGAAAGTCCCATGCTTCTCCCAAAATTAATTCCTCAGAGAT 889
DB 859 AGCTTACTCTGAGCGGATAGTGCCTATCCCTTCAAAACAGCTCAACCTTTACCGAT 918
QY 890 GGTCAATGTTCTGCGGTTGGTGTGTTCTAAGCATCTTCTGCACTACCGTCTCACAAATCC 949
DB 919 TGTTATCATCTCCGCTCTATCATCTCTGATGTTCTTCTTCCAAATATCGGTGCTACATCC 978
```

```
QY 950 TGTGCGTAATGTCATACCCACTGTGGCTTTTATCTGTATATATGAGATTTGGTTGCTTT 1009
DB 979 AGTGGCGGATGCTTATGGAATTGTGGCTAGTATCTGTATCTGTGAAATTTGGTTGCCCTT 1038
QY 1010 ATCTGTGATATCTGATCATGTTCCGGAAGTGGTTTCCAAATCAACCGGAGAGACTTACCTTGA 1069
DB 1039 ATCTGTGCTCTGATCATATTCCTCAAGTGGTACCCTGATTAACCGTGAACATACCTTGA 1098
QY 1070 TAGACTGCTTTTAAAGGTATGACGAGAGGTGAACCTCTCAGTTGCTGTGTTGACAT 1129
DB 1099 CAGGCTTGCATATGATATGATGAGGAGGAGAGGACCATCAAGCTTCTCTCCATGTATGT 1158
QY 1130 ATTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1189
DB 1159 CTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1218
QY 1190 ATCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1249
DB 1219 GTCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1278
QY 1250 AGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1309
DB 1279 TTGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1338
QY 1310 ACCATTTGTGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1369
DB 1339 TCCGTTTGTGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1398
QY 1370 AATTGATATCTGAAAGACAAAGTCCAGCTTCAATTTGTTAAAGACCGCGGCGCATGAA 1429
DB 1399 AATGATATCTGAAAGACAAAGTCCAGCTTCAATTTGTTAAAGAGGCGGCGCATGAA 1458
QY 1430 GAGAGATATGAGAAATTTTAAATCAGATATAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1489
DB 1459 GAGAGATATGAGAAATTTTAAATCAGATATAATGCTGCTGCTGCTGCTGCTGCTGCTG 1518
QY 1490 CCGGAGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1549
DB 1519 ACCTGAGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1578
QY 1550 TCATCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1609
DB 1579 TCACCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1638
QY 1610 TGAGTCTCCCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1669
DB 1639 CGAGTGGCAGGCTTGTCTAGCTCTCTGCGTGAAGAGGCGGAGGATTCAGGATCAAA 1698
QY 1670 GAAGCTGCTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1711
DB 1699 GAAGCTGCTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
```

RESULT 9

US-09-221-013A-3
; Sequence 3, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3

Db 6817 TTGGGATCTATCGAGATCTTTCTTAGCAGACATTTCTCTATCTGATGTTTACCATGGA 6876
QY 2589 CGCCTCAAGTTCTCGGAGAGATTCGCTTACATCAACACCAATTTTACCCATCACTCT 2648
Db 6877 AGGTTGAGACTTTTGGAGAGATCGCTTATATCAACACCATCTCTATCTTACATCC 6936
QY 2649 CTCGCGCTTCTAGTCTATTTGCTATATTTGCTGCTCTATCTGCTCACTGGAAGTTCTATC 2708
Db 6937 ATCCCTCTTATGCGTATTTATTTCTCCCGCTTTTGTCTCATCCGACAGATTCATC 6996
QY 2709 ATGCCAG----- 2715
Db 6997 ATACCCGAGGTTGTAAACCTGACACACTGCTGATTTACTATTTTGAATCCCATTTTGTGA 7056
QY 2716 -----AGATTAGCAACTTGGCCAGTATCTGTTTCAT 2746
Db 7057 ATGCATTTTGTTCATCATCATCTTGTTCAGATAAGCAACTACGCGAGTATTTGTTTCAT 7116
QY 2747 TGGCTCTCTCTTCAATTTTGGCACTGGTATCTTGGATGAGTGGAGTGGTGGTGG 2806
Db 7117 TCTACTCTTCTATCTCAATTTGCTGTGACTGGAATCTTGGAGCTGAGATGGAGCGGTGGAG 7176
QY 2807 CATTGACGAGTGGTGGAGGATGAACAGTCTTGGGTCATTTGGAGGTATCTCTGCACATCT 2866
Db 7177 CATTGAGATTTGGTGGAGGACGAGCAGTTCTGGGTCAATTGGTGGCACATCGGCCATCT 7236
QY 2867 GTTTGCGCTCTTTCAAGGCTCTTGAAGTGGTTCGCGGTATCGACCACTTCACTGT 2926
Db 7237 TTTTGTCTCTTCCAAAGTCTACTTAAAGTCTTGTCTGCTATCGACACCAACTTCCCGT 7296
QY 2927 CACCTCAAGGTAATGACGAGAGAGGAGCTTTTCTGCTGAGCTCTACATGTTCAAGTGGAC 2986
Db 7297 TACATCTAAAGCCACAGAGATGGGATTTTGCAGAACTCTACATCTTCAAAAGGAC 7356
QY 2987 GACGCTCTCTATCCCTCCGACGACCAATTTTGAATCAATTAACATGTTGGTGTCTTGTCTGG 3046
Db 7357 AGCTCTTCTATCTTACCACCAACACCGTCTCTACTTGTGAACCTCATAGGCAITGTGGCTGG 7416
QY 3047 CACCTCTACGCACTCAACAGTGGTTTCAATCATCTGCGGGCGCTCTTTGGGAAGCTCTT 3106
Db 7417 TGCTCTTATGCTGTAACAGTGGTACCACTGCTGAGTGGGCTCTTCTTGGGAACTCTT 7476
QY 3107 CTTTGGCTCTGGGTGATGTTCACTTATACCATTTCTCAAGGCTCTTATCGGACGCA 3166
Db 7477 CTTGCGCTTATGGGTATTTGCCCATCTTACCCCTTCTTGAAGGTCTGTTGGGAGACA 7536
QY 3167 AAACCGCACACCGAGATTTGTCATGCTCTGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCT 3226
Db 7537 AAACCGAACACCAACCATGCTCATTTGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCT 7596
QY 3227 GCTGTGGGTCTGTTGATCCATT 3250
Db 7597 GCTTTGGGTCTGATCCATCCCTT 7620

RESULT 10
US-09-221-013A-4
; Sequence 4, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27

; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 5009
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(5009)
; OTHER INFORMATION: N is A, T, G or C.
US-09-221-013A-4

Query Match 6.1%; Score 219.6; DB 4; Length 5009;
Best Local Similarity 60.4%; Pred. No. 1.7e-49;
Matches 437; Conservative 0; Mismatches 199; Indels 88; Gaps 1;
QY 813 TTACTGATGATGAACCTCGCCAGCCTCTATCTAGAAAGTCCCATGCTCTCTCCAAA 872
Db 2416 TTCAGGATGGATGAAGGAAGAACCTCTCTCAAGAAAGCTACCTATTTGTTCAAGCAGA 2475
QY 873 ATAAATCCCTACAGAAATGGTCAATTTCTGCGGTTGGTTGTTCTAAAGCATCTTCTCGAC 932
Db 2476 ATAAATCCCTACAGATGTTAAATTTCTGTGTCGCTCGCATTTCTTGGTCTTTTCTTCT 2535
QY 933 TACCGTCTCAAAATCCCTGTCGCTAATGATGATGATGATGATGATGATGATGATGATGAT 992
Db 2536 TATAGATTTCTCCATCCAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2595
QY 993 GAGATTTGGTTGCTTTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1052
Db 2596 GAGATATGGTTGCGAGTGTCTTGGATTTCTGATCAATTTCCCAAAATGATGATGATGATGAT 2655
QY 1053 CGGAGAGCTACTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 2656 CGTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2715
QY 1081 -----TAAG 1084
Db 2716 TGCAATATTCATTTTTCATTTTCCAAAAATTTTGTAAACTCTATTTTCTTACATTAAG 2775
QY 1085 GTATGACGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1144
Db 2776 GTACGAG 2835
QY 1145 CGACCCCTTGAAGGAGGACCATCTATGCTCACTGCAACACTGTGCTATCCATTTCTGCTGT 1204
Db 2836 GGATCCGTTGAAG 2895
QY 1205 TGATATCCGTTGGGAG 1264
Db 2896 TGATATCTGTTGGATAG 2955
QY 1265 TTTTGAAGCTGCTCTGATACAGCTGAGTTTCTGATGATGATGATGATGATGATGATGATGAT 1324
Db 2956 ATTTGAAGCTCTCTGATACAGCTGAGTTTCTGATGATGATGATGATGATGATGATGATGAT 3015
QY 1325 GTATGACATTTGAACCCAGAGCTCCGAGTTTCTATTTTGTCCAGAAATTTGATTAAGTAA 1384
Db 3016 GTTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3075
QY 1385 AGACAAAGTCCAGCTTCAATTTGTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1444
Db 3076 GAACAAAGTTCATCTCTGCTTTTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3135
QY 1445 ATTT 1448
Db 3136 TTTT 3139

RESULT 11
US-09-313-294A-6485
; Sequence 6485, Application US/09313294A
; Patent No. 6476212

```

; GENERAL INFORMATION:
; APPLICANT: Lalugudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6485
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700351882H1
; US-09-313-294A-6485

Query Match          5.7%; Score 206.4; DB 4; Length 281;
Best Local Similarity 83.6%; Pred. No. 1.3e-46;
Matches 234; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 2618 CATCAACACCAACATTACCCACTTAACCTCTCTCCCGCTTCTAGTCTATTGTATATGGCC 2677
Db 1 CATCAACACCAACATTACCCGCTCAGGTCCTCCCTCCCTCTCATTTACTGTATCTCTGCC 60

QY 2678 TGTATCTGTCTGCTCACTGGAAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTAT 2737
Db 61 TGGCATCTGCTGCTCACTGCGGGAAGTTTCATCATGCCAGAGATTAGCAACTTGGCTAGTAT 120

QY 2738 CTGGTTCAATGGCGCTCTTCTCTTCAATTTTGGCACTGGTATCTTGAGATGAGTGGAG 2797
Db 121 CTGGTTCAATCTCTCTCTTCAATCTCGATCTTGGCAGCGGTATCTGGAGTGGAGTGGAG 180

QY 2798 TGTGTTGGCAATGACGAGTGTGGAGGAATGAACAGTTCTGGGTCAATGGAGGTATCTC 2857
Db 181 CGCGGTGGCATCGACGAGTGTGGAGGAACGAGCAGTCTGGGTATCGAGGCAATCTC 240

QY 2858 TGCATATCTGTTGGCGTCTTTTCAAGGTCTTTTCAAGGTG 2897
Db 241 CGCCCACTCTTCCGCTCTTCCAGGCGCTCTCTCAAGGTG 280

RESULT 12
US-09-313-294A-4753
; Sequence 4753, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalugudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4753
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348982H1
; NAME/KEY: unsure
; LOCATION: 64, 189
; OTHER INFORMATION: a, t, c, g, or other
; US-09-313-294A-4753

Query Match          5.6%; Score 204; DB 4; Length 300;
Best Local Similarity 82.2%; Pred. No. 6.2e-46;
Matches 245; Conservative 0; Mismatches 52; Indels 1; Gaps 1;

```

```

QY 2716 AGATTAGCAACTTGGCGAGTATCTGGTTATTTGGCTCTTCTTCAATTTTGGCACTG 2775
Db 3 AGATCAGCAACTTGGCGAGCATCTGGTTATCTCTCTCTTCAATCTCGATCTTGGCCACGG 62

QY 2776 GTATCCTTGAGATGAGTGGAGTGTGTGGCAATTCAGAGTGTGGAGGAATGAACAGT 2835
Db 63 GNATCCTTGAGATGAGTGGAGCGGGTGGGCATTCAGCAGTGTGTGGAGGAACGAGCAGT 122

QY 2836 TCTGGGTCAATTTGAGGTATCTCTGCACATCTGTTTCCCGTCTTTTCAAGGTCTTTTGAAG 2895
Db 123 TCTGGGTGATCGGGGGCATCTCCGCGCACCTCTTGGCGGTTCAGGGGCTGTCTCAAGG 182

QY 2896 TGCTTGGCGGTATCGACACCACTTCACTGTCACTCAAGGTATGACGAAGAGCGG 2955
Db 183 TGCTGNGCGCATCGACACCACTTCACTGTCACTTCCCGTCACTTCCAGGGCTCGGAGGAGCGG 242

QY 2956 ACTTGTGCTGAGCTTACATGTTCAAGTGGAGCAGCTTCTCATCCCTCCGACGACCAT 3013
Db 243 ACTTGGGGA-CTGTACATGTTCAAGTGGAGCAGCTTCTTATCCCGCCACCAACCAT 299

RESULT 13
US-09-313-294A-5519
; Sequence 5519, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalugudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5519
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350285H1
; NAME/KEY: unsure
; LOCATION: 6, 40, 54, 65, 67, 200, 221, 230, 238, 261
; OTHER INFORMATION: a, t, c, g, or other
; US-09-313-294A-5519

Query Match          5.5%; Score 198.2; DB 4; Length 291;
Best Local Similarity 81.6%; Pred. No. 2.3e-44;
Matches 235; Conservative 0; Mismatches 52; Indels 1; Gaps 1;

```

```

QY 2428 GAGGCTGGGGTTCAAGTCTATTGTCATGCCCAAGCGCCAGCTTTCAAGGGATCTGCCCCCA 2487
Db 4 GANGCTGGCGATCAATCTACTGTCATGCTTAAGCGACNAGCTTTCAAGGGANCTGCTCCTA 63

QY 2488 T-CAATCTTTGAGATCGTCTGAAACAGTGTGCGGTGGGCTCTCGGTCTCTGTTGAAAT 2546
Db 64 TNCNACCTTTGGGATCGTTTGAATCAAGTGTCTCGGTGGGCTCTTGGTTCCATTTGAAAT 123

QY 2547 CTTTTCAGCGCGCATTTGCCCTTATGGTATGCTACGGAGGGCGCTCAAGTTCTCTGGAG 2606
Db 124 CTTTTCAGCGCGCATTTGCCCATATGGTATGCTATGGAGGTGGCTTAAATTCCTGGAG 183

QY 2607 AGATTGCTTATCAACCAACCACTTATACCACTAACCTCTCTCCCGCTTCTAGTCTAT 2666
Db 184 AGATTGCTTATCANCAACCACTTATACCACTCANATCAATCCNGCTCCTCTGTTAC 243

QY 2667 TGTATATTGCTGCTATCTGTCTGCTCACTGGAAGTTCAATGCA 2714
Db 244 TGCATATTGCCAGCAGTGTCTTCTCACTGGAAGTTCAATGCA 291

```

RESULT 14

US-09-313-294A-4696
; Sequence 4696, Application US/09313294A

; Patent No. 6476212

; GENERAL INFORMATION:

; APPLICANT: Lalgudi, Raghunath V.

; APPLICANT: Ito, Laura Y.

; APPLICANT: Sherman, Bradley K.

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

; FILE REFERENCE: PL-0017 US

; CURRENT APPLICATION NUMBER: US/09/313,294A

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 7600

; SOFTWARE: PERL Program

; SEQ ID NO 4696

; LENGTH: 286

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6476212 700348878H1

; NAME/KEY: unsure

; LOCATION: 91-92, 182, 252

; OTHER INFORMATION: a, t, c, g, or other

US-09-313-294A-4696

Query Match 4.9%; Score 177; DB 4; Length 286;

Best Local Similarity 80.4%; Pred. No. 1.5e-38;

Matches 229; Conservative 0; Mismatches 54; Indels 2; Gaps 2;

Qy 2716 AGATTAGCAACTTGGCCAGTATCTGTTTCATTGGCGCTCTTCTTCAATTTTCGCCACTG 2775

Db 4 AGATCAGCAACTTGGCCAGCACTGTTCACTCCCTCTTCATCTCGATCTTCGCCACGG 63

Qy 2776 GTATCTTGATGATGAGTGGAGTGGTGTGGCAATGACGAGTGGAGGAATGAACAGT 2835

Db 64 GCATCTCGAGATGAGTGGAGCGGGNNCGCATCGAGTGGTGGAGGAACGAGCAGT 123

Qy 2836 TCTGGTCTATGGAGTATCTCTGACATCTGTTTGGCGCTCTTTCAGGGTCTTCTGAGG 2895

Db 124 TCTGGTGTATCGGGGGCATCTCCGGCACCTCTTCGCCGTGTTCAGGGGCTGTCTCAAG 183

Qy 2896 TGCTTGGCGGTATCGACACCACTTCACTGTCACTCAAGGCTAATGACGAAGAAGCG 2955

Db 184 TGCTGGCGGCATCGACACCACTTCACTGTCACTCAAGGCT- CGGACGAGGACGCG 242

Qy 2956 ACTTGTGAGCTCTACATGTTCAAGTGGAGCGAGCTTCTCATCC 3000

Db 243 A-TTCGGGANTGTAACATGTTCAAGTGGAGCGACCTCTCATCC 286

RESULT 15

US-09-313-294A-3993

; Sequence 3993, Application US/09313294A

; Patent No. 6476212

; GENERAL INFORMATION:

; APPLICANT: Lalgudi, Raghunath V.

; APPLICANT: Ito, Laura Y.

; APPLICANT: Sherman, Bradley K.

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

; FILE REFERENCE: PL-0017 US

; CURRENT APPLICATION NUMBER: US/09/313,294A

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 7600

; SOFTWARE: PERL Program

; SEQ ID NO 3993

; LENGTH: 284

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6476212 700347494H1

; NAME/KEY: unsure

; LOCATION: 52, 189, 203, 209, 227, 240, 251, 254, 259, 264, 267, 278-279
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-3993

Query Match 4.0%; Score 146.8; DB 4; Length 284;

Best Local Similarity 74.4%; Pred. No. 2.7e-30;

Matches 203; Conservative 0; Mismatches 68; Indels 2; Gaps 2;

Qy 3069 GGTTACCAATCATGGGGCGCTCTTTGGGAAGCTCTTCTTGGCTTCTTGGGTGATTCGT 3128

Db 1 GGGTACCAGTCTGTGGGTCCGCTCTTCGGCAAGCTCTTCTTGGCTTCTTGGGTGATTCGT 60

Qy 3129 -CACTTATACCCATTCTCAAGGTCTTATGGGAGGCAAAACCGCACCGCAGATTGT 3187

Db 61 ACACCTGTACCCGTTCTCAAGGTCTCATGGTTCGGCAGAACCGCACCCGACCATGT 120

Qy 3188 CATCGT-CTGGGCTGTCTCTCGTCTCTATCTTCTCTCTGTGGTTCGTGTTGATC 3246

Db 121 GGTGTACTGGGCGATCTGTGGGTCGATCTTCTCTCTGTGGTTCGTCATCGATC 180

Qy 3247 CATTCACCTACCGTCTCGCTGGCCCAAAATATCCAAACCTGTGGCATCAACTGCTAGGAAA 3306

Db 181 CGTTCAACNACCGGTAAATTGNCNCGGANTACCGAACGTGTGGCATTAATTGCAAGGAN 240

Qy 3307 GTGGGAGTTTGTAGAGACAGAAAATATAACAGT 3339

Db 241 GTTGAAGTTTNTANAACNAAGANTACAGAT 273

Search completed: August 23, 2004, 05:32:01

Job time : 265 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2004, 05:27:52 ; Search time 1528 Seconds
(without alignments)
11656.672 Million cell updates/sec

Title: US-09-900-237A-29
Perfect score: 3626
Sequence: 1 gcacagagaaccccgctcca.....tcacatttggaggagtttt 3626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3228839 seqs, 2456066551 residues

Total number of hits satisfying chosen parameters: 6457678

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PC7_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PC7US_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	3626	100.0	3626	9	US-09-900-237-29	Sequence 29, Appl
2	2694.4	74.3	3962	17	US-10-437-963-72402	Sequence 72402, A
3	2595	71.6	3799	13	US-10-627-132-21	Sequence 21, Appl
4	2595	71.6	3799	15	US-10-267-459-5	Sequence 5, Appl
5	2595	71.6	3799	15	US-10-209-059-21	Sequence 21, Appl
6	2572.6	70.9	3746	15	US-10-160-719-29	Sequence 29, Appl
7	2572.6	70.9	3746	15	US-10-160-719-49	Sequence 49, Appl
8	2572.6	70.9	3773	13	US-10-627-132-9	Sequence 9, Appl
9	2572.6	70.9	3773	15	US-10-209-059-9	Sequence 9, Appl
10	2572.6	70.9	3773	15	US-10-160-719-5	Sequence 5, Appl
11	2326.2	64.2	3704	13	US-10-627-132-13	Sequence 13, Appl
12	2326.2	64.2	3704	15	US-10-209-059-13	Sequence 13, Appl
13	2326.2	64.2	3704	15	US-10-160-719-57	Sequence 57, Appl
14	2319.8	64.0	3534	13	US-10-425-114-2553	Sequence 2553, Ap

c	15	2316.4	63.9	4264	17	US-10-437-963-39762	Sequence 39762, A
	16	1681.6	46.4	1733	9	US-09-900-237-19	Sequence 19, Appl
c	17	1673.6	46.2	3614	15	US-10-229-193-9	Sequence 9, Appl
	18	1431.4	39.5	3191	13	US-10-424-599-95990	Sequence 95990, A
	19	1408.4	38.8	3725	13	US-10-627-132-45	Sequence 45, Appl
	20	1408.4	38.8	3725	15	US-10-209-059-45	Sequence 45, Appl
	21	1408.4	38.8	3725	15	US-10-160-719-13	Sequence 13, Appl
	22	1408.4	38.8	3725	15	US-10-160-719-21	Sequence 21, Appl
	23	1408.4	38.8	3725	15	US-10-160-719-41	Sequence 41, Appl
	24	1383.6	38.2	3851	16	US-10-393-840-69	Sequence 69, Appl
	25	1378.2	38.0	3723	13	US-10-424-599-106896	Sequence 106896,
	26	1374	37.9	3753	15	US-10-160-719-33	Sequence 33, Appl
	27	1374	37.9	3753	15	US-10-160-719-53	Sequence 53, Appl
	28	1374	37.9	3780	13	US-10-627-132-1	Sequence 1, Appl
	29	1374	37.9	3780	15	US-10-209-059-1	Sequence 1, Appl
	30	1374	37.9	3780	15	US-10-160-719-9	Sequence 9, Appl
	31	1364.4	37.6	3563	9	US-09-900-237-25	Sequence 25, Appl
	32	1343.4	37.0	3603	15	US-10-229-193-5	Sequence 5, Appl
	33	1341.8	37.0	3673	15	US-10-229-193-11	Sequence 11, Appl
	34	1340.2	37.0	3786	9	US-09-900-237-7	Sequence 7, Appl
	35	1340.2	37.0	3813	13	US-10-627-132-17	Sequence 17, Appl
	36	1340.2	37.0	3813	15	US-10-209-059-17	Sequence 17, Appl
	37	1340.2	37.0	3813	15	US-10-160-719-25	Sequence 25, Appl
	38	1340.2	37.0	3813	15	US-10-160-719-45	Sequence 45, Appl
	39	1331.4	36.7	2830	13	US-10-627-132-5	Sequence 5, Appl
	40	1331.4	36.7	2830	15	US-10-267-459-1	Sequence 1, Appl
	41	1331.4	36.7	2830	15	US-10-209-059-5	Sequence 5, Appl
	42	1320.4	36.4	3776	9	US-09-900-237-3	Sequence 3, Appl
	43	1317.2	36.3	3568	13	US-10-627-132-41	Sequence 41, Appl
	44	1317.2	36.3	3568	15	US-10-209-059-41	Sequence 41, Appl
	45	1317.2	36.3	3568	15	US-10-160-719-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-900-237-29
; Sequence 29, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 3626
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-900-237-29

Query Match 100.0%; Score 3626; DB 9; Length 3626;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCACAGGAAACCCCGCTCCAGCTCTGTGTCGGTGGATCGTCTGCCCGCA 60

Db 1 GCACAGGAAACCCCGCTCCAGCTCTGTGTCGGTGGATCGTCTGCCCGCA 60

Qy 61 TGGACGGCAGCGGACGCCCTTGAAGTCGGGAGCAGCGGGACGTTGTCCAGA 120

Db 61 TGGACGGCAGCGGACGCCCTTGAAGTCGGGAGCAGCGGGACGTTGTCCAGA 120

Qy 121 TCTGGCCGACGGCCCTGGGACCACCGTTGGACGGCAGCTTCTACCCGCTCGACGCT 180

121 TCTGCGCGACGCGCTGGGACACAGTTGAGCGCGAGCGTCTTCA CGCGCTCGACGCTCT 180 Db
181 GCGCGTTCGCGGCTCTGCGCGCCCTGTGTACAGAGCA CGAGCGCAAGAGGGGCA CCGAGGCGCT 240 Qy
181 GCGCGTTCGCGGCTCTGCGCGCCCTGTGTACAGAGCA CGAGCGCAAGAGGGGCA CCGAGGCGCT 240 Db
241 GCCTCCAGTCCAGAGCAACCAAGTACAGCGGCCACAGAGGGAGCCAGCGATCCGCGGGGAGG 300 Qy
241 GCCTCCAGTCCAGAGCAACCAAGTACAGCGGCCACAGAGGGAGCCAGCGATCCGCGGGGAGG 300 Db
301 AAGCGGACGACACTGTAGTCCGAGTATGAGTGTGACATTCACCTCCCTGCACTCTGGCACTG 360 Qy
301 AAGCGGACGACACTGTAGTCCGAGTATGAGTGTGACATTCACCTCCCTGCACTCTGGCACTG 360 Db
361 AGGACAGAGAGAGAGATTGCTGACAGAGTGGCGAGCTGGCGGATGAACACCGGGGGCA 420 Qy
361 AGGACAGAGAGAGAGATTGCTGACAGAGTGGCGAGCTGGCGGATGAACACCGGGGGCA 420 Db
421 GTGCGAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCGCTCTCCAAGTATGACA 480 Qy
421 GTGCGAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCGCTCTCCAAGTATGACA 480 Db
481 GTGCGAGATCCCTAGGGGATAGTCCCTTCAGTCAACCAACAGCCAGATCTCAGGAGAAA 540 Qy
481 GTGCGAGATCCCTAGGGGATAGTCCCTTCAGTCAACCAACAGCCAGATCTCAGGAGAAA 540 Db
541 TCCGTGGAGCTTCGCGCTGATCATCAGATGATGCCCTACGGGGAAACATCAGCAGAGTG 600 Qy
541 TCCGTGGAGCTTCGCGCTGATCATCAGATGATGCCCTACGGGGAAACATCAGCAGAGTG 600 Db
601 CTCGGTTTCCCTATGTCAATCATTCACCAATCGTCAAGGGAGTTCTCCGGCAGTATTG 660 Qy
601 CTCGGTTTCCCTATGTCAATCATTCACCAATCGTCAAGGGAGTTCTCCGGCAGTATTG 660 Db
661 GGAATGTTGCTGGAAAGAGAGAGTGTGATGGCTGGAAATGAAGCAGGACAAGGGTCCGA 720 Qy
661 GGAATGTTGCTGGAAAGAGAGAGTGTGATGGCTGGAAATGAAGCAGGACAAGGGTCCGA 720 Db
721 TTCCATGACTAATGGGACAGAGCTTCTCCCTGAAAGTGGCGGAGCTACTGACATCG 780 Qy
721 TTCCATGACTAATGGGACAGAGCTTCTCCCTGAAAGTGGCGGAGCTACTGACATCG 780 Db
781 ATGCATCTACTGAAATACCAATGGAAGACGCTTTACTGAAATGAAGTCAAGCTCGCCAGGCTC 840 Qy
781 ATGCATCTACTGAAATACCAATGGAAGACGCTTTACTGAAATGAAGTCAAGCTCGCCAGGCTC 840 Db
841 TATCTAGAAAAGTCCCAATGCTTCTCCAAAATAAATCCCTACAGAAATGGTCAATTGCTC 900 Qy
841 TATCTAGAAAAGTCCCAATGCTTCTCCAAAATAAATCCCTACAGAAATGGTCAATTGCTC 900 Db
901 TCGGTTGGTGTCTTAAGCATCTTCTGCACTACCGTCTCAGAAATCCCTGTGCGTAATG 960 Qy
901 TCGGTTGGTGTCTTAAGCATCTTCTGCACTACCGTCTCAGAAATCCCTGTGCGTAATG 960 Db
961 CATACCACTGTGCTTTATCTGTATATGTGAGATTGGTTGGCTTTATCTCTGGATAC 1020 Qy
961 CATACCACTGTGCTTTATCTGTATATGTGAGATTGGTTGGCTTTATCTCTGGATAC 1020 Db
1021 TGGATCAGTTCGCGAAGTGGTTTCCAATCAACCGGGAGACCTACTCTGATAGACTGGCTT 1080 Qy
1021 TGGATCAGTTCGCGAAGTGGTTTCCAATCAACCGGGAGACCTACTCTGATAGACTGGCTT 1080 Db
1081 TAAGGTATGACCGAGAGGTGAACCGGCTCAGTTGGCTGTGTTGACATATTTGTCAGTA 1140 Qy
1081 TAAGGTATGACCGAGAGGTGAACCGGCTCAGTTGGCTGTGTTGACATATTTGTCAGTA 1140 Db
1141 CAGTCGACCCCTTGAAGGAGCCACTATCGTCACTGCGCAACACTGTGCTATCCATTCTTG 1200 Qy
1141 CAGTCGACCCCTTGAAGGAGCCACTATCGTCACTGCGCAACACTGTGCTATCCATTCTTG 1200 Db
1201 CTGTTGATATCCCGTGACAAAGGTCTCTGCTAATGATCTGATGACGAGCTTCAATGC 1260 Qy

1201 CTGTTGATATCCCGTGGACAAAGGTCTCTTGCTATGTATCTGATGACGAGCTTCAATGC 1260 Db
1261 TGAATTTTGAAGCATTTGGCTTGAGACTTCAGAGTTTCTAGGAAATGGGTACCAATTTCTGA 1320 Qy
1261 TGAATTTTGAAGCATTTGGCTTGAGACTTCAGAGTTTCTAGGAAATGGGTACCAATTTCTGA 1320 Db
1321 AGAAGTATGACATTTGAACCCAGAGCTCCCGAGTTTACTTTTCCGAGAAATTTGATTACC 1380 Qy
1321 AGAAGTATGACATTTGAACCCAGAGCTCCCGAGTTTACTTTTCCGAGAAATTTGATTACC 1380 Db
1381 TGAAGACAAAGTCCAGCTTCATTTGTTAAAGACCGCGCGGCATGAAGAGAGAAATATG 1440 Qy
1381 TGAAGACAAAGTCCAGCTTCATTTGTTAAAGACCGCGCGGCATGAAGAGAGAAATATG 1440 Db
1441 AAGAAATTTAAATCAGATAAATGCCCTAGTTTCTTAAGGCATTTGAAAGTCCCGAGGAAG 1500 Qy
1441 AAGAAATTTAAATCAGATAAATGCCCTAGTTTCTTAAGGCATTTGAAAGTCCCGAGGAAG 1500 Db
1501 GATGGATCATGCAAGATGGCACACCATGGCCAGGAAACAAATACAGGGATCATCTCTGGAA 1560 Qy
1501 GATGGATCATGCAAGATGGCACACCATGGCCAGGAAACAAATACAGGGATCATCTCTGGAA 1560 Db
1561 TGATTCAGGTTTCTCTGTTGTCACAGTGGTGGCCCTTGATCTAGAGGTAAATGAGCTCCCC 1620 Qy
1561 TGATTCAGGTTTCTCTGTTGTCACAGTGGTGGCCCTTGATCTAGAGGTAAATGAGCTCCCC 1620 Db
1621 GTTTAGTTTATGTTCTCGTGAAGAGCTCTCGGTTCCAGACCAAGAGGCTGGT 1680 Qy
1621 GTTTAGTTTATGTTCTCGTGAAGAGCTCTCGGTTCCAGACCAAGAGGCTGGT 1680 Db
1681 CCATGAATGCCCTTGTGTTCTCAGCTGTCTTAATGAACAATAATGTTGAATC 1740 Qy
1681 CCATGAATGCCCTTGTGTTCTCAGCTGTCTTAATGAACAATAATGTTGAATC 1740 Db
1741 TTGATTTGATCATCTACATCAACAGAGAGCTGTCCAGAGCTATGTTCTCTAA 1800 Qy
1741 TTGATTTGATCATCTACATCAACAGAGAGCTGTCCAGAGAGCTATGTTCTCTAA 1800 Db
1801 TGGATCCAAACCTAGGTCGCAAGTCTGTTATGTGCAAGTTCCTCCAAAGGTTTGAATGGGA 1860 Qy
1801 TGGATCCAAACCTAGGTCGCAAGTCTGTTATGTGCAAGTTCCTCCAAAGGTTTGAATGGGA 1860 Db
1861 TTGATAGGAATGATCGATATGCAAAACAGGAACTGTCTTTTGTATTAATTAACCTGAGGG 1920 Qy
1861 TTGATAGGAATGATCGATATGCAAAACAGGAACTGTCTTTTGTATTAATTAACCTGAGGG 1920 Db
1921 GCCTTGACGCAATCAAGGACAGTTCATGTTGGAACTGTTGTTGTTTCAACAGACAG 1980 Qy
1921 GCCTTGACGCAATCAAGGACAGTTCATGTTGGAACTGTTGTTGTTTCAACAGACAG 1980 Db
1981 CTATCTATGTTATGAGCCCCCAATTAAGGCGAAGAGCCAGGTTTCTTGCGCATCATAT 2040 Qy
1981 CTATCTATGTTATGAGCCCCCAATTAAGGCGAAGAGCCAGGTTTCTTGCGCATCATAT 2040 Db
2041 GTGGGGCAAGAGAGGCAAGCAAGTCAAGAAAGAGGAGCTCAGATAAGAAAAGTCGA 2100 Qy
2041 GTGGGGCAAGAGAGGCAAGCAAGTCAAGAAAGAGGAGCTCAGATAAGAAAAGTCGA 2100 Db
2101 ACAAGCATGTGGACAGTCTGTTCCAGTATTCATCTCGAAGACATAGAGGAGGCTGTTG 2160 Qy
2101 ACAAGCATGTGGACAGTCTGTTCCAGTATTCATCTCGAAGACATAGAGGAGGCTGTTG 2160 Db
2161 AAGGTGCTGGGTTTGTATGATGAGAAATCAGTTCCTCATGTCTCAATCAGCTTAGAGAAGA 2220 Qy
2161 AAGGTGCTGGGTTTGTATGATGAGAAATCAGTTCCTCATGTCTCAATCAGCTTAGAGAAGA 2220 Db
2221 GATTTGGCCAGTCAGCAGCAATTTGTTGCCCTCCACTCTGATGGAATATGGTGGTTCCTC 2280 Qy
2221 GATTTGGCCAGTCAGCAGCAATTTGTTGCCCTCCACTCTGATGGAATATGGTGGTTCCTC 2280 Db
2281 AGTCTCTCACTCCAGATCTCTTTGAAAGAGCTATCCATGTCATAGTTGCGGCTATG 2340 Qy
2281 AGTCTCTCACTCCAGATCTCTTTGAAAGAGCTATCCATGTCATAGTTGCGGCTATG 2340 Db

```
QY 2341 AGGACAAAGTCTGAATGGGAACTGAGATTGGTTGGATCTATGGATCTGTACAGAGATA 2400
Db 2341 AGGACAAAGTCTGAATGGGAACTGAGATTGGTTGGATCTATGGATCTGTACAGAGATA 2400
QY 2401 TTCTTACTGGATTCAAGATGCAGCAGAGGGTGGGGTTTCAGTCTATTGATGCCCAAGC 2460
Db 2401 TTCTTACTGGATTCAAGATGCAGCAGAGGGTGGGGTTTCAGTCTATTGATGCCCAAGC 2460
QY 2461 GGCAGCTTTCAGGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTCTGC 2520
Db 2461 GGCAGCTTTCAGGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTCTGC 2520
QY 2521 GGTGGGCTCTCGTCTCTGTTGAAATCTTTTCAGCCGCGATTGCCCTTATGGTATGGCT 2580
Db 2521 GGTGGGCTCTCGTCTCTGTTGAAATCTTTTCAGCCGCGATTGCCCTTATGGTATGGCT 2580
QY 2581 ACGAGGGGCGCTCAAGTTCTGGAGAGATTGCTTACATCAACACCAATTTACCCAC 2640
Db 2581 ACGAGGGGCGCTCAAGTTCTGGAGAGATTGCTTACATCAACACCAATTTACCCAC 2640
QY 2641 TAACTCTCTCCGCTTCTAGTCTATTGATTAATGCTGCTGCTATCTGCTGCTCACTGGAA 2700
Db 2641 TAACTCTCTCCGCTTCTAGTCTATTGATTAATGCTGCTGCTATCTGCTGCTCACTGGAA 2700
QY 2701 AGTTTCATCATGCGAGATTAGCACTTCGCCAGTATCTGTTTCAATGCGCTCTTCTTT 2760
Db 2701 AGTTTCATCATGCGAGATTAGCACTTCGCCAGTATCTGTTTCAATGCGCTCTTCTTT 2760
QY 2761 CAATTTTGGCCACTGGTATCTTGGAGATGAGTGGTGGTTCAGTACGAGTGT 2820
Db 2761 CAATTTTGGCCACTGGTATCTTGGAGATGAGTGGTGGTTCAGTACGAGTGT 2820
QY 2821 GGAGGAATGAAAGTCTGGGTCATTGGAGGATCTCTGCACATCTGTTTGGCGTCTTTC 2880
Db 2821 GGAGGAATGAAAGTCTGGGTCATTGGAGGATCTCTGCACATCTGTTTGGCGTCTTTC 2880
QY 2881 AGGGTCTCTGAAGTGTCTGCGGTATCGACCAACTTCACGTCTCACTCAAAAGGCTA 2940
Db 2881 AGGGTCTCTGAAGTGTCTGCGGTATCGACCAACTTCACGTCTCACTCAAAAGGCTA 2940
QY 2941 ATGACGAAGAGGCGACTTTTGTGAGCTCTACATGTTCAAGTGGACGAGCTTCTCATCC 3000
Db 2941 ATGACGAAGAGGCGACTTTTGTGAGCTCTACATGTTCAAGTGGACGAGCTTCTCATCC 3000
QY 3001 CTCGACGACCAATTTGATCATTAACATGGTTGGTGTCTGCTGCTGCTGCTCTCAAGCCA 3060
Db 3001 CTCGACGACCAATTTGATCATTAACATGGTTGGTGTCTGCTGCTGCTCTCAAGCCA 3060
QY 3061 TCAACAGTGGTTTACCAATCATGGGGGCGCTCTTTGGGAAGCTCTTCTTGGCTTCTGGG 3120
Db 3061 TCAACAGTGGTTTACCAATCATGGGGGCGCTCTTTGGGAAGCTCTTCTTGGCTTCTGGG 3120
QY 3121 TGATTTGTCATTTACCCATTCTTCAAGGGTCTTATGGGAGGCAAAACCGCACACCA 3180
Db 3121 TGATTTGTCATTTACCCATTCTTCAAGGGTCTTATGGGAGGCAAAACCGCACACCA 3180
QY 3181 CGATTTGTCATCTGCTGGGTCCTCTGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3240
Db 3181 CGATTTGTCATCTGCTGGGTCCTCTGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3240
QY 3241 TTGATCCATTCACTACCCGCTCTCGTGGGCCCAATATCAAACTGTGGCATCAACTGCT 3300
Db 3241 TTGATCCATTCACTACCCGCTCTCGTGGGCCCAATATCAAACTGTGGCATCAACTGCT 3300
QY 3301 AGGAAAGTGGGAGTTTGTAGACAGAAATATAACAGTATCGAGCAACACCGCGCA 3360
Db 3301 AGGAAAGTGGGAGTTTGTAGACAGAAATATAACAGTATCGAGCAACACCGCGCA 3360
QY 3361 GCCAGAGAAATATTATTTGGGTTGTGAATTTACGTTTGAAGAGTTGTCAAAATTTG 3420
Db 3361 GCCAGAGAAATATTATTTGGGTTGTGAATTTACGTTTGAAGAGTTGTCAAAATTTG 3420
```

```
QY 3421 AGAAACACATTGTAAATAGATGTAATAGACCATCTACGTTTTCATGAGGTTAAGCTC 3480
Db 3421 AGAAACACATTGTAAATAGATGTAATAGACCATCTACGTTTTCATGAGGTTAAGCTC 3480
QY 3481 TTCTTTTGTGAAACAAAGGAATCTCATTGGTAAACCTATAGGAATTTTCTTATAGGCA 3540
Db 3481 TTCTTTTGTGAAACAAAGGAATCTCATTGGTAAACCTATAGGAATTTTCTTATAGGCA 3540
QY 3541 CTTTGATTGTAGGAATGACCTATGAAATGTTGTATTTTATTATATATAAAATTATTC 3600
Db 3541 CTTTGATTGTAGGAATGACCTATGAAATGTTGTATTTTATTATATATAAAATTATTC 3600
QY 3601 TGTCTCTCACATTTCGAGGAGTTT 3626
Db 3601 TGTCTCTCACATTTCGAGGAGTTT 3626
```

RESULT 2

```
US-10-437-963-72402
; Sequence 72402, Application US/10437963
; Publication NO. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 72402
; LENGTH: 3962
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MBT4530_72784C.1
US-10-437-963-72402
```

Query Match 74.3%; Score 2694.4; DB 17; Length 3962;

Best Local Similarity 88.9%; Pred. No. 0;

Matches 2924; Conservative 0; Mismatches 361; Indels 3; Gaps 1;

```
QY 55 GGGCCATGACGGCGACGGGACGCCCTGAACTCGGGAGGACCGGGCCGGGACGTGT 114
Db 368 GGGCCATGACGGCGACGCCCTGAACTCGGGAGGACCGGGAGCGGGAGCGGT 427
QY 115 GCCAGATCTGGCCGACCGCCCTGGGCGACACGCTTGGACGGCGACGCTTTCACCGCTGCG 174
Db 428 GCCAGATCTGGCGGACCGCGCTGGGCGACGACGGCGGAGGCGACGTTTCGCCGCTGCG 487
QY 175 AGCTCTCGCGTTCCCGTCTGCGCCCTGCTACGACGACGAGCGCAAGAGGAGCAACC 234
Db 488 AGCTCTCGCGTTCCCGTCTGCGCCCTGCTACGAGTACGAGCGCAAGGAGGAGCAACC 547
QY 235 AGGCTGCTCTCAGTGCAGACCACTGATGCGGATGTTGAGTGAATCACTTACCTGATCTG 294
Db 548 AGGCTTGGCCCCAGTGCAGACCAAGTACAAGCGGCCCAAGGGGAGCGCGGATCCGCTG 607
QY 295 GGGAGAAAGGCGACACACTGATGCGGATGTTGAGTGAATCACTTACCTGATCTG 354
Db 608 GGGAGAAAGGCGAGGATGATGCTGATGATGTCAGTACTACACTACCTGATCTG 667
QY 355 GCATCTAGGACACAGAGCAGAGATTTGCTGACGATCGCAGCTGGCGCATGAACACCG 414
Db 668 GCAGTGGCGACACAGAGCAGAGATTTGCTGATGATGTCGAGTTCGCGCATGAATCTG 727
QY 415 GGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACGTGGCGAGATCGGCCCTCTCCAAGT 474
```

Db 728 GGGGTGGTGGAGCTCGGCGCTCCCAAGTATGACAGTGGCGAGATCGGCTCACCAAGT 787
Qy 475 ATGACAGTGAGAGATCCCTTAGGGGATACGTCCTCTCAGTCAACAAGCCAGATGTCAG 534
Db 788 ATGACAGTGGGAGATCCCTCGGGATACATCCCTCAGTCACTAATAGCCAGATCTCGG 847
Qy 535 GAGAAATCCCTGAGACTTCGCTGATCATCATGATGTCCTTACCGGGAACATCAGCA 594
Db 848 GAGAAATCCCTGAGACTTCGCTGATCATCATGATGTCCTTACCGGGAACATGCGCA 907
Qy 595 GAGTGTCTCGCTTCCCTATGTGAATCAATCAACAATCCGTCAAGGGAGTCTCCGGCA 654
Db 908 AGCGTCTCAATCCCTATGTGNAACATTCACCAATCATCAAGGGAGTCTCTGGCA 967
Qy 655 GTATGGGAATGTGCTCGTGAAGAGAGAGAGTGTGATGGCTGGAAATGAAGCAGACAGG 714
Db 968 GCATTGGAATGTGCTCGAAGAGAGAGTGTGATGGCTGGAAATGAAGCAGACAGG 1027
Qy 715 GTCGATTCCTCATGACTAATGAGGACAGCATGTCCTCTGAGGTTCGGGAGCTACTG 774
Db 1028 GAGCAATTCCTCATGACCGGACAGCATGTCCTCTGAGGTTCGGGAGTGTGCTG 1087
Qy 775 ACATCGATGCATCTACTGAATACAAATGGAAGACGCTTTACTGAATGATCAAACTGCC 834
Db 1088 ATATCGATGCATCCACTGATTAACAATATGGAAGATGCTTACTGAATGATGAATCGCC 1147
Qy 835 AGCCTCTATAGAAAGTCCCATTCGCTTCTCCAAATTAATATCCCTCAGAAATGTCAC 894
Db 1148 AGCCTCTCTAGAAAAGTTCCTCTTCAATCAGAAATAATCCCTCAGAAATGTCAC 1207
Qy 895 TTGTTCTCGCGTGTGTTCTAAGCATCTTCGCACTACCGTCTCACAATTCCTCTGTC 954
Db 1208 TTGTTCTCGCGTGTGTTCTAAGCATTTTCTTCACTACCGTATTCAGATTCCTGTC 1267
Qy 955 GTAATGCATACCCACTGTGGCTTTATCTGTATATGATGAGATTTGTTTGTCTTATCCT 1014
Db 1268 GCAATGCTATCCGCTTCGCTTTATCTGTATATGATGAGATTTGTTTGTCTTCT 1327
Qy 1015 GATATCGATCAGTTCGGAAGTGTTCGAATCAACCGGAGACCTACCTGATAGAC 1074
Db 1328 GATATTCGATCAGTTCGGAAGTGTTCGAATCAACCGGAGACCTACCTGATAGAC 1387
Qy 1075 TGGCTTTAAGTATGACCGAAGGTGAACCGTCTCAGTTCGCTGCTGTTGACATATTG 1134
Db 1388 TGGCATTCAGGTATGACAGAGAGTGAGCACTCAGTTCGCTGCTGCTGTTGACATTTTG 1447
Qy 1135 TCAGTACAGTCGACCCCTTGAAGAGGCCACTATCGTCACTGCGCAACACTGTGCTATCCA 1194
Db 1448 TCAGTACAGTCGACCCCATGAAGGAACCTCTCTTGTACTGCGCAATACCGTGTGTCCA 1507
Qy 1195 TTCTTGTGTTGATATCCGTGGAAGGTCTCTTGTATGATCTGATGACGAGCTT 1254
Db 1508 TTCTTGTGTTGATATCCGTGGAAGGTCTCTTGTATGATCTGATGATGATGATGATG 1567
Qy 1255 CAATGCTGACTTTTGAAGCATTCGCTGAGACTTCAGAGTTCCTAGGAAATGGGTACCAT 1314
Db 1568 CAATGCTGACTTTTGAAGCATTCGCTGAGACTTCAGAGTTCCTAGGAAATGGGTACCAT 1627
Qy 1315 TTGTGAAGAGATGATGACATTAAGACCCAGAGCTCCCGAGTTCCTGTTTGTGCGAGAAATG 1374
Db 1628 TCGTTAAGAGATTAACATTCAGGCAAGAGCTCTGATGATCTCTTCTCCAGAAATCG 1687
Qy 1375 ATTAACCTGAAAGCAAGAGTCCAGCTTCATTTGTTAAGACCGCGGCGCATGAAGAGAG 1434
Db 1688 ATTAACCTGAAAGCAAGAGTCCAGCTTCATTTGTTAAGACCGCGGCGCATGAAGAGAG 1747
Qy 1435 AATATGAAGATTTAAATCAGGATTAATGCGCTAGTTTCTAAGGCAATGAAAGTCCCG 1494
Db 1748 AATATGAAGATTTAAATCAGGATTAATGCGCTAGTTTCTAAGGCAATGAAAGTCCCG 1807
Qy 1495 AGGAAGGATGATCATGCAAGTGGCACCAATGCGGAGGAAACATATCAGGAGATCATC 1554

1808 AGGAGGATGATCATGCAAGATGGCACACCATGGCCAGGAAACAATACTAGGACCATC 1867
Qy 1555 CTGGAATGATTCAGGTTTCCCTTGGTCAAGTGGTGGCTTGATACTAGAGGTAATGAGC 1614
Db 1868 CTGGAATGATTCAGGTTTCCCTTGGTCAAGTGGTGGCTTGATACCGAGGTAATGAGC 1927
Qy 1615 TCCCCCGTTTATGTTTATGTTCTCTGTGAAAGCGTCTCTGGGTTCCAGACCAACAAGG 1674
Db 1928 TTCCCGCGCTGTCTACGATATCTCTGTGAGAGCGTCTCTGGGTTTACGACCAACAAGG 1987
Qy 1675 CTGCTGCATGAATATCCCTTGGTCTCTGCTCTCAGTGTCTCTTAATAATGACATATG 1734
Db 1988 CTGCTGCATGAATATCTCTGTCTCTGCTCTCAGTGTCTCTTACCAATGACAGTACATG 2047
Qy 1735 TGAATCTGATCTGATCACTACATCAACAACAGCAGGCTGTCCGAGAGCTATGTGCT 1794
Db 2048 TGAATCTGATCTGATCACTACATCAACAACAGCAGGCTGTCCGGAAGCTATGTGCT 2107
Qy 1795 TCCTAATGATCCAAACCTAGGTCCGCAAGTCTGTTTATGTGCAGTTCCTCCACAAGGTTG 1854
Db 2108 TCCTAATGATCCAAACCTAGGAGGAGTGTCTGTTATGTTCAGTTCCTCCACAAGGTTG 2167
Qy 1855 ATGCGATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1914
Db 2168 ATGCGTATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2227
Qy 1915 TGAGGGGCTTTCAGCGCATTCAGGACAGTTCATGTTGAGGAACTGGTGTGTTTCAACA 1974
Db 2228 TGAGGGGCTTTCAGCGCATTCAGGACAGTTCATGTTGAGGAACTGGTGTGTTTCAACA 2287
Qy 1975 GAACAGCTATCTATGTTTATGAGCCCCCAATTA---AGGCGAAGAGCGAGTTCCTGG 2031
Db 2288 GAACCTGCTATATGTTTATGAGCCCCCAATTAAGCAGAGAGAGAGGAGGAACTTCCTGT 2347
Qy 2032 CATCATGATGTTGGGGCAAGAGAGCAGCAGTCAAGAAAGAGAGCTCAGATGAGA 2091
Db 2348 CATCATGATGTTGGGGCAAGAGAGCAGCAGTCAAGAAAGAGAGTTCGAGACAAGA 2407
Qy 2092 AAAAGTCGAACAAGCATGTCAGATGTCCTCCAGTATTCATCTCGAAGACATAGAGG 2151
Db 2408 AGAAGTCAACAAGCAGCAGTGCCTGTCGAGTTTCAATCTTGAAGATATAGAGG 2467
Qy 2152 AGGCTGTGAAGGTGCTGGGTTTGAATGAGAAATCAGTTCATCTCAATAGAGCT 2211
Db 2468 AGGCTGTGAAGGTGCTGGATTCGATGATGAGAAATCACTTCTTATGCTCAATAGAGCT 2527
Qy 2212 TAGAAGAGATTTGGCAGTCAAGCAGATTTGTTGCTCCACTCTGATGGAATATGCTG 2271
Db 2528 TGGAGAAGAGATTTGGCAGTCTGAGCGTTCCTGCTCCACTCTCATGGAATATGCTG 2587
Qy 2272 GTGTTCTCAGTCTCCACTCCAGAACTCTTTTGAAGAGAGCTATCCATGTCATAAGTT 2331
Db 2588 GTGTTCTCAATCTGCAACCCAGAACTCTTTTGAAGAGAGCTATCCATGATGATAGTT 2647
Qy 2332 GTGCTATGAGGACAACTCTGAAATGAGGAACTGAGATTTGGTGGATCTATGGAATGCTCA 2391
Db 2648 GTGCTATGAGGACAAAGCCGAAATGAGGAACTGAGATTTGGGAGTCTACGTTTCCGTA 2707
Qy 2392 CAGAGATATTTTACTGGAATTCAGATGCAAGAGAGGCTGCGTTTCACTCTATGCA 2451
Db 2708 CAGAGATATTTTACTGGAATTCAGATGCAAGAGAGGCTGCGTTTCACTCTATGCA 2767
Qy 2452 TGCCCAAGCGCCAGCTTTCAAGGGAATCTGCCCCCATCAATCTTTCAGATGCTGTAACC 2511
Db 2768 TGCCCAAGCGCCAGCTTTCAAGGGAATCTGCTCTCTATCAATCTTTCAGATGCTGTAACC 2827
Qy 2512 AAGTGTGCGGTGGGCTCTCGGTTCTGTTGAAATCTTTTTCAGCGCGCATGCCCCCTAT 2571
Db 2828 AAGTGTGCGGTGGGCTCTCGGTTCTGTTGAAATCTTTTTCAGTTCGCGCATGCCCCCTAT 2887
Qy 2572 GGTATGCTACGAGGCGGCTCAAGTTCTTGGAGAGATTCGCTTACATCAACACCACCA 2631
Db 2888 GGTACCGCTATGGAGGAGCGCTTAAAGTTCTTGGAGAGATTTGCTTACATCAACACCACCA 2947

QY	936	CGTCTCACAATCCTGTGCGTAATGTCATACCCACTGTGGCTTTTATCTGTTATATCTGAG	995
Db	1114	CGGATCACAATCCTGTGCGTAATGCATACCCACTGTGGCTTCTATCTGTTATATGTGAG	1173
QY	996	ATTGTGGTTTGCTTTATCCTGTGATACTGGATCAGTTCGGAAAGTGGTTTCCAATCAACCG	1055
Db	1174	ATCTGGTTTGCTCTTTTCTGTGATATTGGATCAGTTCAGTTCCTCAAAGTGGTTTCCAATCAACCGC	1233
QY	1056	GAGACCTACCTTGATAGACTGGCTTTAAGGTATGACCGGAAAGTGAACCGTCTCAGTTG	1115
Db	1234	GAGACTTACCTTGATAGACTCGCATTAAGGTATGACCGGAAAGTGAACCGTCTCAGTTG	1293
QY	1116	GCTGCTGTGACATATTGTCAGTACAGTTCGACCCCTTGAAGGAGCCACCTATCGTCAC	1175
Db	1294	GCTGCTGTTGACATTTTGTGACGTACTGTGACCCCAATGAAGGAGCCTCCTCTTGTCAC	1353
QY	1176	GCCAACTGTGCTATCCATTCCTTGTGTTGATATCCCGTGGACAAGTCTCTTGCTAT	1235
Db	1354	GCCAACTACCGTGCTATCCATTCCTGCTGTGGACTATCCTGTGGATAAGTCTCTTGCTAT	1413
QY	1236	GTATCTGATGACGAGCTTCAATGCTCACTTTTTCAGCCATTGGCTGAGACTTCAGAGTTT	1295
Db	1414	GTATCTGATGATGAGCTGCTATGCTGCATTTGATGCATGAGCTGAGACTTCAGAGTTT	1473
QY	1296	GCTAGGAATGGGTACCATTTGTGAAGAAGTATGACATTGAACCCAGAGCTCCCGAGTTT	1355
Db	1474	GCTAGGAATGGGTGCCATTTGTTAAGAAGTACAACATTGAACCTAGAGCTCCTGAATGG	1533
QY	1356	TACTTTTCGAGAAATTTGATTAATCTGAAGACAAGTCCAGCCTTCATTTGTTAAAGAC	1415
Db	1534	TACTTCTCCAGAAATTTGATTAATCTGAAGGACAAAGTGCCCTTCATTTGTTAAAGAC	1593
QY	1416	CGCGGGCCATGAAGAGAGATATCAAGAAATTAATAATCAGGATAAATGCCAGTTTCT	1475
Db	1594	CGCGGGCCATGAAGAGAGATATGAAGAAATCAAAATTAGGTPAATGGCTTGTTGCT	1653
QY	1476	AAGCATTTGAAAGTCCCGAGGAAGGATGATCATGCAAGATGCCACCATGGCCAGGA	1535
Db	1654	AAGCACAAANAAGTCCCTGAGGAAGGATGGATCATGCAAGATGCCACCATGGCCAGGA	1713
QY	1536	AACAACTACAGGGATCATCTCTGGAATGATTCAGGTTTTTCTTGTCACAGTGGTGGCTT	1595
Db	1714	AACAATACAGGGACCATCTCTGGAATGATTCAGGTTTTTCTTGCTCAGTGGTGGTCTT	1773
QY	1596	GATCTGAGGGTAAATGAGCTCCCGGTTTGTGTTTATGTCTCTGTAAGGCGTCTGGG	1655
Db	1774	GATCTGAGGGTAAATGAGCTACCCCGTTTGTGCTATGTTTCTGTAAGAACGTCCTGGA	1833
QY	1656	TTCCAGCACCAAGAAGGCTGGTGCCATGAATGCCCTTGTTCGTGTCTCAGCTGTGCTT	1715
Db	1834	TTCCAGCATCAAGAAGCTGGTGCCATGAATGCTCTTGTCGCGTCTCAGCTGTGCTT	1893
QY	1716	ACTAATGACATACATGTTGAAATCTTGATTTGTGATCATCATCATCAACAGCAAGCT	1775
Db	1894	ACCAATGACATAATGTTGAAATCTTGATTTGTGATCACTACATCAACAGTAAAGCT	1953
QY	1776	GTCCGAGAGCTATGTGCTTCTTAATGATCCAAACCTTAGTGTCCGAACTGTTATGTG	1835
Db	1954	CTCAGGAGAGCTATGTGCTTCTTAATGATCCTAACCTAGGAAGAGTGTCGTGATGTT	2013
QY	1836	CAGTTTCCCAAGAGTTTGTATGGGATTCATAGGAATGATCGATATGCAACAGGAACACT	1895
Db	2014	CAGTTTCCCAAGAGTTTCGATGGTATTGATAGGAATGATCGATATGCAACAGGAACACC	2073
QY	1896	GTCTTTTGTGATTAATTAATGAGGGGCTTGACGGCATTTCAAGGACCAAGTTTATGTGGGA	1955
Db	2074	GTGTTTTTGTGATTAATTAATGAGAGGCTTCGATGGCATCCAAAGGACCAAGTTTATGTGGGC	2133
QY	1956	ACTGTTGTGTTTTTCAACAGAACAGCTATCTATGTTTATGAGCCCCCAATTAAGGCGAAG	2015
Db	2134	ACTGGCTGTGTTTTTCAACAGAACAGCTCTATATGTTTATGAGCCCCCAATTAAGCAAAG	2193
QY	2016	AAGCCAGGTTTTCTTGGCATCTACTATGTGGGGGCAAGAAGAGGCAAGCAAGTCAAGAAA	2075

[illegible]

Db 3271 GCGAAGCTCTTCTTCCGCTCTGCGGTGATGCTTACCTGTATACCCGTTCTCCTCAAGGCTCTC 3330
QY 3156 ATGGCAGGCAAAACCGCACACCGAGATTGTCTATCGTCTGGCTGTCTCTCTCGTCTT 3215
Db 3331 ATGGGTGCGCAACCGCACCCCGACCATCGTGTGTCTGGCGATCTCTGCGCTCG 3390
QY 3216 ATCTTCTCTCTGCTGCTGCGTGTGATGCTGATCATCTACTACCCGTTCTCGTGGCCCAAT 3275
Db 3391 ATCTTCTCTGCTGCTGCGTGTGATGCTGATCATCTACTACCCGTTCTCGTGGCCCGAT 3450
QY 3276 ATCCAAACCTGTGGCATCAACTGCTAGGAAAGTGGGA-GTTGTAGACACAGAA 3329
Db 3451 ACTCGAAGCTGTGGCATCAACTGCTAGGAGTGAAGGTTGTAGAAACACAGA 3505

RESULT 4
US-10-267-459-5
; Sequence 5, Application US/10267459
; Publication No. US20030150014M1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864RC
; CURRENT APPLICATION NUMBER: US/10/267,459
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3799
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (238)...(3477)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(3799)
; OTHER INFORMATION: n = A,T,C or G
US-10-267-459-5

Query Match 71.6%; Score 2595; DB 15; Length 3799;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 2871; Conservative 0; Mismatches 420; Indels 4; Gaps 2;

QY 36 CGGTTTGATCGCTCTGCGGCCATGACCGCGACCGCGACGCCCTGAAGTCCGGGAGG 95
Db 214 CGGTTCTGGGCGAGCTCGCTGCCATGAGGCGGACCGCGACCGCGGAGG 273
QY 96 CACGGGCGGGAGCGTGTGCGAGATCTGCGCGACCGCTGGGACACACGTTGGAACGC 155
Db 274 CGCGGGGAGGCGAGGTGTGCGAGATCTGCGCGGATGCGTGGGCATACCGCGGAGGGA 333
QY 156 GACGCTTTCACCGCTGCGAGCTGTGCGGCTTCCCGGCTCTCGCGCCCTGCTACGAGCAC 215
Db 334 GACGCTTTCACCGCTGCGAGCTGTGCGGCTTCCCGGCTGTGCGCGCCCTGCTACGAGTAC 393
QY 216 GAGCGCAAGGAGGCGACCCAGGCTGCTTCCAGTSCAAGACCAAGTACAAGCGCCACAGA 275
Db 394 GAGCGCAAGGAGGCGACCAAGCGTGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 453
QY 276 GGGAGCCAGCATCCGCGGGGAGGAGCGACGACGACCTGATCGCGATGATGTAGTAC 335
Db 454 GGGAGTCCAGCGATCCGAGGGGAGGAGGAGACGATCTGATCGCGATGATGTAGTAC 513
QY 336 TTCAACTACCCTGCTGCTGCGCTGAGGACCGAGGACGAGAGATTGCTGACAGGATGCGC 395

Db 514 TTCAACTACCTGCTGATCTGGCAATGACGACCGAAGAGAGATTGCTGACAGGATGCGC 573
QY 396 AGCTGGCGCATGAACACACCGGGGCGAGTGGCAATGTGTGGCCACCCCAAGTATGACAGTGGC 455
Db 574 AGCTGGCGCATGAATGCTGGGGGCGAGCGGGGATGTGGCGCGCCCAAGTATGACAGTGGT 633
QY 456 GAGATCGGCTCTCCAGTATGACAGTGGAGAGATCCCTAGGGGATAGCTCCCTTCAGTC 515
Db 634 GAGATCGGCTCTTACCAAGTACGAGTGGTGGATCCCTCGGGGATATCATCCCGTCACTC 693
QY 516 ACCAACCGCAGATGTCTAGGAGAAATCCCTGGAGCTTCGCTGATCATCATCATGATGTC 575
Db 694 ACTAACCGCAGATTTCCGGGAGAAATCCCTGGTGTCTCCCTGACCATCATGATGTC 753
QY 576 CCTACGGGGAACATCAGCAGCTGCTCCGTTTCCCTATGTGAATCATTCACCAATCCG 635
Db 754 CCTACTGGGAACATTTGGCAGGCGGCCCAATTTCCCTATATGAATCATTCATCAATCCG 813
QY 636 TCAAGGGAGTTTCTCGGCGAGTATTTGGGAATGTTGGCTGGAAAGAGAGAGATTGATGGCTGG 695
Db 814 TCGAGGGAAATTTCTCTGCTAGCGTTGGGAATGTTGCCCTGGAAAGAGAGGTTGATGGCTGG 873
QY 696 AAAATGAAGCAGGACCAAGGTTGCCGATTCCTCATGACTAATGGACAAGCATTCCTCCCTCT 755
Db 874 AAAATGAAGCAGGACCAAGGAAACAAATTCCTCATGCAAGTATGGCAAGCATTCCTCCCTCT 933
QY 756 GAAAGTGGGCGAGCTACTGACATCATCATCATCTACTGAATCAACATGAGAGAGAGCTTTA 815
Db 934 GAGGCGCGGGGTGTGTGTGATTTGATGATCATCAACTGATTACAACTGAGAGAGTGGCTTA 993
QY 816 CTGAATGATGAACCTCGCCAGCCTCTATCTAGAAAGTCCCATGCTTCCTCCAAATA 875
Db 994 TTAACCGATGAACCTCGCCAGCCTCTATCTAGAAAGTTCACACTTCCTTCCTCCAGGATA 1053
QY 876 AATCCCTACAGATGGTTCATTTGTTCTGCGGTGTTGTTCTTAAGCATCTTCTCGCAGTAC 935
Db 1054 AATCCATACAGAGTGGTTCATTTGCTGACGATGATTTCTAGAGCATCTTCTGACATAC 1113
QY 936 CGTCTCAAAATCCTGTGCGTAATGCAATCCACCTGTTGGCTTTTATCTGTTATATGTAG 995
Db 1114 CGGATCAAAATCCTGTGCGTAATGCAATCCACCTGTTGGCTTCTATCTGTTATATGTAG 1173
QY 996 ATTTGGTTGCTTTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1055
Db 1174 ATCTGGTTGCTCTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1233
QY 1056 GAGACCTACCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1115
Db 1234 GAGACTTACCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1293
QY 1116 GCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1175
Db 1294 GCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1353
QY 1176 GCGAAGCTGTGCTATCCATTTCTGCTGTTGATTTATCCGTTGGAGAGGCTCTCTGCTAT 1235
Db 1354 GCGAATACCGTGTATCCATTTCTGCTGTTGATTTATCCGTTGGAGAGGCTCTCTGCTAT 1413
QY 1236 GTATCTGATGACGAGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1295
Db 1414 GTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1473
QY 1296 GCTAGGAAATGGGTACCATTTGTGAAGAGTATGACATTTGAACCCAGAGCTCCCGAGTTT 1355
Db 1474 GCTAGAAAATGGGTGCCATTTGTTAAGAGTATCAACATTTGAACCTAGAGCTCCCTGAATGG 1533
QY 1356 TACTTTTCCAGAAAATTTGATCTGAAAGACAAAGTCCAGCTCTCATTTGTTTAAAGAC 1415
Db 1534 TACTTCTCCAGAAAATTTGATCTGAAAGACAAAGTCCAGCTCCCTCATTTGTTTAAAGAC 1593
QY 1416 CGCGGCGCATGAGAGAGATATGAAGAAATTTTAAATCAGATAAATGCCCTAGTTTCT 1475

Db 1594 CGCGGGCCATGAAGAGAGATATGAAGAATTCAAATTTAGGTAATGGCCCTTGTTGCT 1653
Qy 1476 AAGGCATTGAAGTCCCAGAGAGATGATCATCAAGATGCGACACCATGGCCAGCA 1535
Db 1654 AAGGCACAAAAATGCTTGGAGAGAGATGATCATCAAGATGCGACACCATGGCCAGCA 1713
Qy 1536 AACAAATACCAAGGATCATCTCGAATGATTCAGGTTTTCCTTGGTCACAGTGGTGCCCTT 1595
Db 1714 AACAAATACCAAGGACCATCTCGAATGATTCAGGTTTTCCTTGGTCACAGTGGTGCTT 1773
Qy 1596 GATACAGGAGTAATGAGCTCCCCGTTTATGTTTATGTTCTCGTCAAAAGCGTCTGGG 1655
Db 1774 GATACAGGTAATGAGCTACCCGTTTGGTCTATGTTCTCGTCAAAAGCGTCTGGG 1833
Qy 1656 TTCCAGCACCAAGAGCTGTGCGATGAATGCGCTTGTGTTGTTCTCAGCTGTCCTT 1715
Db 1834 TTCCAGCATCACAAAGAGCTGTGCGATGAATGCTTGTGCGCGTCTCAGCTGTCCT 1893
Qy 1716 ACTAATGGCAATACATGTTGAATCTTGATTTGTGATCACTACATCAACACAGCAAGGCT 1775
Db 1894 ACCAATGGCAATACATGTTGAATCTTGATTTGTGATCACTACATCAACACAGTAAGCT 1953
Qy 1776 GTCCGAGAGCTATGCTTCTTAATGATCCAAACCTAGTCCGCAAGTCTCTTTATGTT 1835
Db 1954 CTGAGGAAAGCTATGCTTCTTATGGATCTTAACCTAGGAAGAGTGTCTGCTATGTT 2013
Qy 1836 CAGTCCCAAAAGTTTGTATGGATTTGATGGATTTAGGAATGATCGATCAACACAGCAACT 1895
Db 2014 CAGTTTCCCAGAGGTTGATGGATTTGATGGATTTGATGGATTTGATGGATTTGATGGG 2073
Qy 1896 GTCTTTTTCATATTAACTTTAGGGGCTTTGACGGCATTTCAAGGACAGTTTATGTTGGA 1955
Db 2074 GTGTTTTTCATATTAACTTTAGAGGCTTTGATGGCATTTCAAGGACAGTTTATGTTG 2133
Qy 1956 ACTGGTTGTTTTCACAGAAACAGCTATCTATGTTATGAGCCCCCAATTAAGGCGAAG 2015
Db 2134 ACTGGCTGTGTTTTCACAGAAACAGCTCTATATGTTATGAGCCCCCAATTAAGCAAAAG 2193
Qy 2016 AAGCAGGTTCTTGGCATCACTATGTTGGGGCAAGAAAGCAAGCAAGTCAAGAGAA 2075
Db 2194 AAGGGTGGTTTCTGTCATCACTATGTTGGGAGGAAGGAAGGAAGCAATCAAGAA - 2252
Qy 2076 AAGAGCTCAGATAAGAAAAGTCCAAACAGCATGTGGACAGTTCTGTTCAGTATTCAAT 2135
Db 2253 --GGGCTCAGACAAGAAAAGTCCAAACAGCATGTGGACAGTTCTGTGCGCATTTCAAT 2310
Qy 2136 CTCGAAGACATAGAGAGGGTGTGAAGTGTGGGTTTGAATGATGAGAAATCAGTTCTC 2195
Db 2311 CTTGAAGATATAGAGGAGGGAGTTGAAGGCGCTGGATTTGATGATGAGAAATCACTTCT 2370
Qy 2196 ATGTCCTCAATGAGCTTAGAGAGAGATTTGGCCAGTCAGCAGCATTTGTCCTCCACT 2255
Db 2371 ATGTCCTCAATGAGCTTAGAGAGAGATTTGGCCAAATCTGACGCTTTGTTGGCTCCACT 2430
Qy 2256 CTGATGGAATATGGTGGTGTCTCAGTCCCTCCAGATCCAGAAATCTCTTTTGAAGAAGCT 2315
Db 2431 CTGATGGAATATGGTGGTGTCTCAGTCCCTCCAGATCCAGAAATCTCTTTGAAAGAAGCT 2490
Qy 2316 ATCCATGCTAATAGTTGTGGCTATGAGGACAAAGTCTGAATGGGGAATGAGATTGGTGG 2375
Db 2491 ATCCATGCTAATAGTTGTGGCTACGAGGACAAAGTGAATGGGGAATGAGATTGGTGG 2550
Qy 2376 ATCTATGGATCTCTCAGAGAAATATTTCTTACTGGATTCAAGATGACGCAAGAGGCTGG 2435
Db 2551 ATCTATGGTCTCTGACGGAAGATATTTCTCAGTGGGTTCAAGATGACGCAAGAGGCTGG 2610
Qy 2436 CGTTCAGTCTATTGTCATGCCCAAGCGCCAGCTTTTCAAGGATCTGCCCCCAATCAATCTT 2495
Db 2611 CGGTGCAATCTACTGTCATGCCCAAGCGCCAGCTTTTCAAGGATCTGCCCCCAATCAATCTC 2670
Qy 2496 TCAGATGCTGACCAAGTGTGCTGGTGGCTCTGGTCTGTTGAAATCTTTTCAGC 2555
Db 2671 TCAGACCGTCTGAACCAAGGTGCTCCGGTGGGCTCTGGGTTTCAGTGGAAATCTCTTTTCAGC 2730

Qy 2556 CGGCATTGGCCCTTATGGTATGGCTACGAGGCGCCTCAAGTTCTCTGGAGATTCGCT 2615
Db 2731 CGGCATTGGCCCTTATGGTATGGCTACGAGGCGCCTCAAGTTCTCTGGAGATTCGCT 2790
Qy 2616 TACATCAACACCAACCAATTTACCCACTAACCTCTCTCCCGCTTCTAGTCTATTGTATATTG 2675
Db 2791 TACATCAACACCAACCAATTTACCCACTAACCTCTCTCCCGCTTCTAGTCTATTGTATATTG 2850
Qy 2676 CTTGCTATCTGTCTGTCTAGTCTGGAAGTTTATCATGCGAGATTAAGCAACTTGGCCAGT 2735
Db 2851 CTTGCGCATCTGTCTGTCTGCGGGAAGTTTATCATATCCAGAGATCAAGCAACTTCGCTAGT 2910
Qy 2736 ATCTGTTTCAATTCGCTCTTCTTCAATTTTCGCACTGGTATCTCTGAGATGAGGTGG 2795
Db 2911 ATCTGTTTCAATTCGCTCTTCTTCAATTCGCACTGGTATCTCTGAGATGAGGTGG 2970
Qy 2796 AGTGGTGTGGCAATGACGAGTGGTGGAGAAATGAACAGTTCTGGGTCTATTGAGGTATC 2855
Db 2971 AGCGGCTGGGCATCGACGAGTGGTGGAGAAACGAGCAGTTCTGGGTCTATCGGAGCATC 3030
Qy 2856 TCTGCACTATCTGTTTGGCCGCTTTCAGGCTCTTCTGAAGTCTTCTCCGATTCGACAC 2915
Db 3031 TCGGCCACCTCTTTCGCGCTTTCAGGCGCTTCTCAAGTGTCTTCCGCGCATCGACAC 3090
Qy 2916 AACTTCACCTGTCACTTCAAGGCTAATGACGAAGAGGCACTTTGCTGAGCTTCTACATG 2975
Db 3091 AACTTCACCGTCACCTTCAAGGCTCGGATGAAGACGCGACTTTCGCGAGCTGTACATG 3150
Qy 2976 TTCAAGTGAAGCAGCTTCTATCTCTCCGACGACCAATTTTGTATCAATTAACATGGTTGGT 3035
Db 3151 TTCAAGTGAAGCAGCATTCTGATCCCGCCACACCATCTGATCATCAACCTGGTCGGC 3210
Qy 3036 GTGCTTGTGGCAGCTTCTACGCCATCAACAGTGGTTTACCAATCATGGGCGCGCTCTTT 3095
Db 3211 GTTGTTCGCGCATCTTCTACGCCATCAACAGGCGGTACAGTGGTGGGCTCCGCTTC 3270
Qy 3096 GGAAGCTCTTCTTTCGCTTCTGGGTGATTGTTTCACTTATACCATTTCTTCAAGGCTCTT 3155
Db 3271 GGAAGCTCTTCTTTCGCTTCTGGGTGATTGTTTCACTTATACCATTTCTTCAAGGCTCTC 3330
Qy 3156 ATGGGAGGAAACACGACACCGCATGTTGTCATGTTGGCTGTCTGCTCTCTCTCTCT 3215
Db 3331 ATGGGTGGGAGAAACCGCACCCCGCATCTGTTGTTGTTCTGGGCGATCTCTGCTGGGCTG 3390
Qy 3216 ATCTTCTCTTGTGTGGTTCGTTGTTGATTCATTCACCTGCTCCGCTCTGCTGCCAAAT 3275
Db 3391 ATCTTCTCTTGTGTGGTTCGATTCGATTCGTTTCAACACCGCTCTGCTGGCCGAT 3450
Qy 3276 ATCCAAACCTGTGGCATCACTCTAGGAAAGTGGGA-GTTTGTAGAGACAGAAA 3329
Db 3451 ACTCGAAGCTGTGGCATCACTCTAGGAGGTGGAAAGTTCGTAGAAACAGAGA 3505

RESULT 5

US-10-209-059-21
; Sequence 21, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864P2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 3799
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 3757..3775, 3777, 3782
; OTHER INFORMATION: n = A,T,C or G
US-10-209-059-21

Query Match
Best Local Similarity 71.1%; Score 2595; DB 15; Length 3799;
Matches 2871; Conservative 0; Mismatches 420; Indels 4; Gaps 2;

QY 36 CGGGTTGGATCGCTCTGCCGCGCCATGAGCGGGAGCGGAGCCCTGAAGTCCGGGAGG 95
DB 214 CGGTTCTGGCGGAGCTCGCTCCCATGGAGGGCGACGCGGCTGAAGTCGGGGAGG 273
QY 96 CACGGGGCGGGAGACGTGTGCCAGATCTCGCGGACGGCTGGGACACACGTTGGACGGC 155
DB 274 CGCGGGGAGGCGAGGTGTGCCAGATCTCGGGGATGGCGTGGGCTACTAGCGGGAGGA 333
QY 156 GACGCTTCAACCGCTGCGACGCTGCGGCTTCCCGGTCTGCCGCCCTCGCTACGAGCAC 215
DB 334 GACGCTTCAACCGCTGCGACGCTGCGGCTTCCCGGTCTGCCGCCCTCGCTACGAGTAC 393
QY 216 GAGCCCAAGGAGGACCCAGCGCTGCTCCAGTGCAGACCAAGTACGAGCCGACAGA 275
DB 394 GAGCCCAAGGAGGACCAACAGCGTGTCCCGGCTGCAAAACAAAGTACAAAGCGCCACAAG 453
QY 276 GGGAGCCAGCGATCCGCGGGGAGGAGCGGACGACACTGATGCCGATGATGATGATGAC 335
DB 454 GGGAGTCCAGCGATCCGAGGGGAGGAGGAGACGATCTGATGCCGATGATGATGATGATG 513
QY 336 TTCAACTACCTTGCATCTGGCACTGAGGACCAAGAGCAGAAGATTCTGACAGGATGCGC 395
DB 514 TTCAACTACCTTGCATCTGGCAATGACGACCAAGAGCAGAAGATTCTGACAGGATGCGC 573
QY 396 AGCTGGCGCATGAACACCGGGGGAGTGCATGTTGGCCACCCAGATGACAGTGGC 455
DB 574 AGCTGGCGCATGAATGCTGGGGGCGAGCGGGGATGTTGGCGGCCCAAGTATGACAGTGT 633
QY 456 GAGATCGGCTCTCAAGTATCACAGTGGAGAGATCCCTAGGGGATAGTCCCTTCAGTC 515
DB 634 GAGATCGGCTTACCAAGTACACAGTGGTGGATCCCTCGGGGATACATCCCTCGAGTC 693
QY 516 ACCAAGCCAGATGTCAGGAGAAATCCCTGGAGCTTCGCTGATCATCATGATGATGCC 575
DB 694 ACTAAGCCAGATTTCCGGAGAAATCCCTGGTGTCTCCCTTGACCATATGATGCT 753
QY 576 CCTACGGGAGACATCAGCAGAGTCTCGGTTCCCTATGTAATCATTCACCAATCGG 635
DB 754 CCTACTGGGAACATTTGGCAGGCGCGCCCATTTCCCTATATGAATCATTCACAAATCGG 813
QY 636 TCAAGGGAGTTCTCCGGCAGTATGGGAATGTCCTGGAAGAGAGAGTTCATCGCTGG 695
DB 814 TCGAGGGAATCTCTGTTAGCTTGGGATGTTGCTGGAAGAGAGGTTGATGGCTGG 873
QY 696 AAAATGAAGCAGGACAAGGGTGCATTCCTAGTAAATGGAAGAGATGCTCCCTCT 755
DB 874 AAAATGAAGCAGGACAAGGGAAATTCCTATGACGAATGGCACAAGCATTCCTCCCTCT 933
QY 756 GAAGTTCGGGAGCTACTGACATCATGATCTACTGATACAACTGGAAGAGCGCTTTA 815
DB 934 GAGGCGCGGGGTGTTGGTGATATTGATGATCAACTGATTAACATGGAAGATGCCTTA 993
QY 816 CTGAATGATGAACCTCGCCAGCTCTATCTAGAAAGTCCCATGCTTCCTCCAAATA 875
DB 994 TTAACAGATGAACCTCGCCAGCTCTATCTAGAAAGTTCACATTCCTTCCTCCAGGATA 1053
QY 876 AATCCCTACAGAAATGGTCAATGTTCTGGGGTGGTGTTCCTAAGCATCTTCCTGCACTAC 935
DB 2134 ACTGGCTGTGTTTTCAACAGAAACAGCTCTATATGGTTATGAGCCCCCANNATTAGCAAAAG 2193
```

```
DB 1054 AATCCATACAGGATGGTCAATGTGTGTACGATGATGTTGTTCTAAGCATCTTCTTGCACTAC 1113
QY 936 CGTCTCACAAATCCTGTGTGCGTAATGCATACCACCTGTGGCTTTTATCTGTATATGTGAG 995
DB 1114 CGGATCACAAATCCTGTGTGCGTAATGCATACCACCTGTGGCTTCTATCTGTATATGTGAG 1173
QY 996 ATTTGGTTTGTCTTTATCCTGGATACCTGGATCAGTTCCTCCGAAGTGGTTTCCAAATCAACCG 1055
DB 1174 ATCTGGTTTGTCTTTCTTGGATATTTGGATCAGTTCCTCAAGTGGTTTCCAAATCAACCGC 1233
QY 1056 GAGACCTACTCTGATAGACTGGCTTTAAGGTATGACCGAGAAGGTGAACCGTCTCAGTTG 1115
DB 1234 GAGACTTACTTTGATAGACTCGCAITTAAGGTATGACCGGGAAGGTGAGCACTCAGTTG 1293
QY 1116 GCTGCTTTGACATATTTGTCTAGTACAGTCGACCCCTTGAAGGAGCCACTATCGTCACT 1175
DB 1294 GCTGCTTTGACATATTTGTCTAGTACGTCGACCCCAATGAAGGAGCCTCTCTGTGTCACT 1353
QY 1176 GCCAACACTGTCTATCCATCTCTGCTGTGATATCCGCTGGACAAAGCTCTCTGCTAT 1235
DB 1354 GCCAATACCGTCTATCCATCTCTGCTGTGGACTATCTGTGGATAAGGTCTCTTGCTAT 1413
QY 1236 GTATCTGATGACGAGCTTCAATGTGCTTTTGAAGCAATGGCTGAGACTTTCAGAGTTT 1295
DB 1414 GTATCTGATGATGAGCTGTATGCTGACATTTGATGACACTAGCTGAGACTTTCAGAGTTT 1473
QY 1296 GCTAGAAATGGGTACCATTTGTGAAGAGTATGACATTTGAACCCAGAGCTCCCGAGTTT 1355
DB 1474 GCTAGAAATGGGTGCCATTTGTTAAGAGTACAAACATTTGAACCTAGAGCTCTGAATGG 1533
QY 1356 TACTTTTCCGAAATTTGATTTACCTGGAAGCAAAAGTCCAGCCTTCATTTCTTAAAGAC 1415
DB 1534 TACTTCTCCGAAATTTGATTTACTTTGAAGGCAAAAGTGCACCCCTTCATTTGTTAAAGAC 1593
QY 1416 CGCCGGGGCCATCAAGAGAGATATCAAGAAATTTAAATCAGGATAAATGCCCTAGTTTCT 1475
DB 1594 CGCCGGGGCCATGAAGAGAGATATCAAGAAATTTCAAAATTAGGGTAAATGGCCTTGTGT 1653
QY 1476 AAGGCAATTAAGTCCCGAGGAGAGATGATGATGCAAGATGGCAACCATGGCCAGGA 1535
DB 1654 AAGGCACAAAAAGTCCCTGAGGAAGATGATGATGCAAGATGGCAACCATGGCCAGGA 1713
QY 1536 AACATACCGGGATCATCTCGGAATGATTCAGGTTTCTTGTGTCAGTGGTGGCCTT 1595
DB 1714 AACATACCGGGACCATCTCGGAATGATTCAGGTTTCTTGTGTCAGTGGTGGCCTT 1773
QY 1596 GATACTGAGGGTAATGAGCTCCCGCTTTAGTATTTATGTCTCGTGAAGAGCTCTCGG 1655
DB 1774 GATACTGAGGGTAATGAGCTACCCGCTTTGGTCTATGTTTCTCGTGAAGAAAGCTCTCGGA 1833
QY 1656 TTCAGACACCAAGAGAGCTGGTGCATGAATGCCCTTGTGTCGTCTCAGCTGTCTT 1715
DB 1834 TTCAGACATCAAGAAAGCTGGTGCATGAATGCTCTTGTCTCGGCTCTCAGCTGTCTT 1893
QY 1716 ACTAATGACAAATACATGTTGAATCTTGATTTGATGATCACTACATCAACCAAGCAAGCT 1775
DB 1894 ACCAATGACAAATACATGTTGAATCTTGATTTGATGATGATGATGATGATGATGATGAT 1953
QY 1776 GTCGAGAGAGCTATGTGCTTCTTAATGATCAAAACCTAGGTCCGCAAGTCTGTTATGTG 1835
DB 1954 CTCAGGGAAGCTATGTGCTTCTTATGATCCTTAACCTAGGAGGAGTGTCTGTATGTT 2013
QY 1836 CAGTTCCCAAAAGTTTGTATGGGATGATAGGAATGATGATGATGATGATGATGATGATGAT 1895
DB 2014 CAGTTCCCAAGAGTTTCGATGGTATTTGATAGGAATGATGATGATGATGATGATGATGAT 2073
QY 1896 GTCCTTTTGTATATTAATCTTGAAGGCTTCAAGGCAATTCAGGACCAAGTATGATGATGAT 1955
DB 2074 GTGTTTTTCGATATTAATCTTGAAGGCTTTCGATGGCATCCCAAGGACCAAGTTTATGTGGC 2133
QY 1956 ACTGGTGTGTTTTCAACAGAAACAGCTATCTATGTTGTTATGAGCCCCCANNATTAGCAAAAG 2015
DB 2134 ACTGGTGTGTTTTCAACAGAAACAGCTCTATATGGTTATGAGCCCCCANNATTAGCAAAAG 2193
```

Qy	2016	AAGCAGGTTCTTGGCATCACTATCTGGGGCAAGAGCAGCAGTCAAGAAA	2075
Db	2194	AAAGGTGGTTCTTGTTCATCACTATCTGGTGGCAGGAAGGAAACAAAGAA-	2252
Qy	2076	AGGAGCTCAGATAAGAAAAGTCCGAAACAGCATGTGGACAGTTCCTGTTCCAGTATTCAT	2135
Db	2253	--GGGCTCAGACNAGAAAAGTACAGAGCATGTGGACAGTTCCTGTCAGTATTCAT	2310
Qy	2136	CTCGAAGACATAGAGAGGTTGTAAGGTGCTGGGTTTGATGATGAGAAATCAGTCTC	2195
Db	2311	CTTGAAGATATAGAGAGGAGTTGAAGGCGCTGGATTTGATGATGAGAAATCACTCTT	2370
Qy	2196	ATGCTCAATGAGCTTAGAGAGAGATTTGGCCAGTACGACGATTTGTCCTCCACT	2255
Db	2371	ATGCTCAATGAGCTTGGAGAGAGATTTGGCCAACTCGACGTTTGTGGTCCACT	2430
Qy	2256	CTGATGAATATGTTGGTTCCTCACTCCACTCCAGAACTCTCTTTGAAAAGAGCT	2315
Db	2431	CTGATGAATATGTTGGTTCCTCACTCTCGACTCCAGAACTCTCTTCTGAAAAGAGCT	2490
Qy	2316	ATCCATGTCTAAGTTGTGCTATGAGGACAAAGTCTGAATGGGAACTGATGTTGGTGG	2375
Db	2491	ATCCATGTCTAAGTTGTGCTACGAGGACAAAGTGAATGGGAACTGATGTTGGTGG	2550
Qy	2376	ATCTATGGATCTGTACAGAGATATTTCTTACTGGATTCAAGATGCAAGCAAGAGGCTGG	2435
Db	2551	ATCTATGGTCTGTGACGAGAGATATTTCTCACTGGGTTCAAGATGCAAGCAAGAGGCTGG	2610
Qy	2436	CGTTCAGTCTATTGTCATGCCCAAGCGCCAGCTTTTCAAGGATCTGCCCCCAATCAATCTT	2495
Db	2611	CGTTCAGTCTATTGTCATGCCCTAAGCGCGCCCTTCAAGGATCGGCTCCCAATCAATCTC	2670
Qy	2496	TCAGATGCTTGAAACAAAGTCTGCGTGGGCTCTGGTTCGTTGAAATCTTTTCAGC	2555
Db	2671	TCAGACGCTGTGAACCAAGTGTCTCGGTGGGCTCTCGGTTCAAGTGAATCTTTTCAGC	2730
Qy	2556	CGCATTTGCCCTTATGTTATGCTACGAGGGCGCTCAAGTTCTCTGGAGAGATTCCGCT	2615
Db	2731	CGCATTTGCCCTTATGTTATGCTACGAGGGCGCTCAAGTTCTCTGGAGAGATTCCGCT	2790
Qy	2616	TACATCAACACCAATTTACCACTAACCTCTCTCCCGTCTTCAGTCTATTGATATG	2675
Db	2791	TACATCAACACCAATCTACCCGCTCACGCTCCCTCCGCTCTCAATTTACTGTATCCTG	2850
Qy	2676	CCTGCTATCTGTGCTCACTGGAAGTTTCATATGCGAGGATTCAGCACTTGGCCAGT	2735
Db	2851	CCTGCCATCTGCTGCTCACGGGAAAGTTTCATATCCAGAGATTCAGCACTTTCGCTAGT	2910
Qy	2736	ATCTGGTTTCAATGGCTCTTCTTCAATTTTCGCCACTGCTATCTTGAAGATGAGGTGG	2795
Db	2911	ATCTGGTTTCAATCTCTTCAATCTCGATCTTCCGACGGGTATCTTGGAGATGAGGTGG	2970
Qy	2796	AGTGGTGTGGCATTTGACGAGTGTGAGAGATGAACAGTCTTGGGTCTATTGAGGTATC	2855
Db	2971	AGCGGCTGGGCATCGACGAGTGTGGAGGAAACGAGCAGTTCTGGGTCTATCGGAGGATC	3030
Qy	2856	TCTGCATCTGTTTGGCTTTTCAGGTTCTTCAAGTGTCTTGAAGTGTCTGCGGTATCGACCC	2915
Db	3031	TCCGCCACCTCTTTCGGCGCTTCTCCAGGGCTCTTCAAGGTGCTTTCGGGATCGACCC	3090
Qy	2916	AACTTCACTGTCACTCAAGGCTTAATGAACGAAGGCGACTTTTGTGAGCTCTACATG	2975
Db	3091	AACTTCACTGTCACTCAAGGCTTCCGATGAAGACGCGACTTTCGGGAGCTGTACATG	3150
Qy	2976	TTCAAGTGGACGAGCTTCTCATCCCTCCGACGACCAATTTTGTATCATTAACATGGTGGT	3035
Db	3151	TTCAAGTGGACGACACTTCTGATCCCGCCCAACCACTCTGTATCATCAACCTGGTCGGC	3210
Qy	3036	GTGTTGCTGGCCTCTTACGCCATCAACAGTGTGTTACCAATCATGGGGCGCTCTTT	3095
Db	3211	GTGTTGCTGGCCTCTTACGCCATCAACAGCGGGTACCAAGTGTGGGTCCGCTCTTC	3270
Qy	3096	GGGAGCTCTCTTTTGCCTTCTGGGTGATTTGTTCACTTATACCATTCCTCAAGGGTCTT	3155
Db	3271	GGCAAGCTCTTCTTTCGCTTCTGGGTGATCGTTCACTGTACCGTTCTCTCAAGGGTCTC	3330
Qy	3156	ATGGGAGGCAAAACCGCACACCGACGATGTGCTGTGGCTGTCTCTCTGCTTCT	3215
Db	3331	ATGGGTGGCAGAAACCGCACACCGACCATGTGTGTTGTTCTGGGCGATCTCTGCTGGGGTGG	3390
Qy	3216	ATCTTCTCTTCTGCTGGGTTCGTTGATTCATTCACCTACCGCTCTCGCTGGCCCAAAAT	3275
Db	3391	ATCTTCTCTTCTGCTGGGTTCGATTCATTCACCTACCGCTCTCGCTGGCCCGGAT	3450
Qy	3276	ATCCAAACCTGTGGCATCAACTGCTAGGAAAGTGGAA-GTTTGTAGAGACAGAAA	3329
Db	3451	ACTCGAACGTGTGGCATCAACTGCTAGGAGGTGGAAGGTTTGTAGAAAACAGAGA	3505
RESULT 6			
US-10-160-719-29			
; Sequence 29, Application US/10160719			
; Publication No. US20030167528A1			
; GENERAL INFORMATION:			
; APPLICANT: Kanwarpal S. Dhugga			
; APPLICANT: Timothy G. Helentjaris			
; APPLICANT: Benjamin A. Bowen			
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses			
; TITLE OF INVENTION: Thereof			
; FILE REFERENCE: 0864			
; CURRENT APPLICATION NUMBER: US/10/160,719			
; CURRENT FILING DATE: 2002-06-03			
; PRIOR APPLICATION NUMBER: US/09/371,383A			
; PRIOR FILING DATE: 1999-08-06			
; PRIOR APPLICATION NUMBER: US 60/096,822			
; PRIOR FILING DATE: 1998-08-17			
; NUMBER OF SEQ ID NOS: 60			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 29			
; LENGTH: 3746			
; TYPE: DNA			
; ORGANISM: Zea mays			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (321)...(3551)			
US-10-160-719-29			
Query Match			
Best Local Similarity 70.9%; Score 2572.6; DB 15; Length 3746;			
Matches 2865; Conservative 0; Mismatches 439; Indels 9; Gaps 2;			
Qy	4	CGAGGAACCCCGCTCCAGCTCTCTGTCGGTGGGGTTGGATCGCTCTGCCGCCATGG	63
Db	265	CCAGGCCCCAGGCTCCAGGCCAGTCCCTCGACGTTTCTCGGGAGCTCGCTTGCATGG	324
Qy	64	ACGGGACGCGGACCGCCCTGAAGTCCGGGAGGACCGGGCCGGGACGTGTGCCAGATCT	123
Db	325	AGGCGGACGCGGACCGCGTGAAGTCCGGGAGGCGCGGTGGCGGACAGTGTGCCAGATCT	384
Qy	124	GGCGGACGCGCTGGGACCCACCGTTGGACCGGACGCTTTCACCGCTCGACGCTTGGCC	183
Db	385	GGCGGACGCGGCTGGGACCCACCGGAGGGGACGCTTTCGCCGCTTGCACGCTTGGC	444
Qy	184	GCTTTCGCGTCTGCCGCCCTTGTCTACGACGACGAGCGCAAGGAGGGACCCAGCGCTGCC	243
Db	445	GGTTCGCGTGTGCCGCCCTTGTCTACGAGTACGAGCGCAAGGACGCGCACGCGCGTGGC	504
Qy	244	TCAGTTCGACGACCAAGTACAAAGCGGACAGAGGAGCCGAGCGATCCGGGGAGGAG	303
Db	505	CCCAGTTCGACGACCAAGTACAAAGCGGACAGAGGAGCCGCGGATCCGTTGGGAGGAG	564
Qy	304	GGCAGGACACTGATCGCGGATGATGGTAGTGACTTCAACTACCTTGCATCTGGCACTGAGG	363
Db	565	GAGACGACACTGATGCCGA-----TAGGACCTTCAATTACCTTGCATCTGGCAATGAGG	618

QY 364 ACCAGAAGCAGAGATTTGCTGACAGATGCGCAGCTGGCGATGAACACCGGGGCGAGTG 423
Db 619 ACCAGAAGCAGAGATTTGCGGACAGATGCGCAGCTGGCGATGAACGTTGGGGGCGG 678
QY 424 GCAATGTTGGCCACCCCAAGTAGTACAGTGGGAGATCGGCCCTCTCCAAAGTAGACAGTG 483
Db 679 GGGATGTTGGTCGCCCCCAAGTAGTACAGTGGGAGATCGGGCTTACCAAGTAGACAGTG 738
QY 484 GAGAGATCCCTAGGGGATAGTCCCTTCAGTCAACAAAGCCAGATGTCAGGAGAAATCC 543
Db 739 GCGAGATTCCTCGGGGATACATCCCATCAGTCAGTCAACAGCCAGATCTCAGGAGAAATCC 798
QY 544 CTGGAGCTTCGCTGATCATCATCATGATGTCCTCCCTACGGGGAACATCAGCAGCCTGTC 603
Db 799 CTGGTCTTCCTCCCTGACCATCATGATGATGTCCTCCCACTGGGAACATGCGAAGCGTGTC 858
QY 604 CTTTCCCTATGTGAATCATTCACCAAAATCCGTCAAGGGAGTTCCTCCGCGCATTTGGGA 663
Db 859 CATTTCCCTATGTGAACCATTCGCCAAATCCGTCAAGGGAGTTCCTGGTASCATTTGGA 918
QY 664 ATGTTGCCCTGGAAGAGAGAGATGTATGGCTGGAAATGAAGCAGACAGAGGTGGATTC 723
Db 919 ATGTTGCCCTGGAAGAGAGAGGTGTATGGCTGGAAATGAAGCAGACAGAGGGAGATTC 978
QY 724 CCATGACTAAATGGGCAAGCATTTGCTCCCTCTGAAAGTTCGGGCGAGTACTGACATCGATG 783
Db 979 CCATGACGATGGCAGACAGCATTTGCTCCCTCTGAGGTCGGGGTGTGGTGATATTGATG 1038
QY 784 CATCTACTGAATPACACATGGAAGCGCTTTTACTGAAATGATGAAACTCGCCAGCCTCTAT 843
Db 1039 CATCAACTGATTACAACTGGAAGATGCTTATTGAAAGCAAGCAAGCGAGCATTC 1098
QY 844 CTAGAAAAGTCCCATTTGCTTCCCTGCAAAAATAATCCCTAAGATGTCATGTTCTGCG 903
Db 1099 CTAGAAAAGTCCCATTTGCTTCCCTCAGGATTAATCCATACAGGATGTCATTTGCTGCG 1158
QY 904 GGTGCTGTTGTTCTAAGCATTTCTTGCACTACCGTCTCACAATCTGTGCGTAATGCAT 963
Db 1159 GATTCATTTGTTCTAAGCATTTCTTGCACTACCGTATCAAAATCCTGTGCGCAATGCAT 1218
QY 964 ACCCATGCTGGCTTTTATCTGTTATATGTGATTTGTTGCTTTATCTCGTGGACTGG 1023
Db 1219 ACCCATATGGCTTCTATCTGTTATATGTGAGATCTGGTTGCTCTTTCGTGGATTTGG 1278
QY 1024 ATCAGTTCGCGAAGTGGTTTCCAATCAACCGGGAGACCTTCACTGATGATGCTGTTAA 1083
Db 1279 ATCAGTTCCTCAAGTGGTTTCCAATCAACCGGGAGACGTACCTTGATAGGCTGGCATTA 1338
QY 1084 GGTATGACCGGAGAGGTGAACCGTCTCAGTTGGCTGCTGTTGACATATTTGTCAGTACAG 1143
Db 1339 GGTATGACCGGAGAGGTGAGCCATCTCAGTTGGCTGCTGTTGACATTTTCGTCAGTACAG 1398
QY 1144 TCGACCCCTTGAAGAGCCACCTATCGTCACTGCCAACACTGTGCTATCCATTTCTGCTG 1203
Db 1399 TCGACCCCAATGAAGAGCCCTCTCTTGTCACTGCCAATAGCGTGTATCCATTTCTTCTG 1458
QY 1204 TGTATTATCCGTGACAAAGTCTCTTGTGATGATGATCTGATGACCGGAGCTTCAATGCTGA 1263
Db 1459 TGGATTTACCCCTGTGGATGAAGTCTCTGCTATGATGATCTGATGAGAGTGGATGCTGA 1518
QY 1264 CTTTTGACGCTATGGCTGAGACTTCAGAGTTTGGCTAGGAAATGGGTACCATTTTGTGAAGA 1323
Db 1519 CATTTGATGCACTAGCTGAGACTTCAGACTTCAGAGTTTGGCTAGGAAATGGGTACCATTTGTTGAAGA 1578
QY 1324 AGTATGACATTGAACCCAGAGCTCCCGAGTTTACTTTTGGCAGAAAATGATTAACCTGA 1383
Db 1579 AGTACAACATTGAACCTAGAGCTCTGAATGATGATCTTCTCCAGAAAATGATTAACCTGA 1638
QY 1384 AAGACAAAGTCCAGCTTCATTTGTTAAAGCCCGCGGCGCATGAAGAGAGAAATATGAAG 1443
Db 1639 AGGACAAAGTCCAGCTTCATTTGTTAAAGCCCGCGGCGCATGAAGAGAGAAATATGAAG 1698

QY 1444 AATTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCATTTGAAAGTCCCGGAGGAGAT 1503
Db 1699 AATTCAAAGTTAGGGTAAATGGCCCTTGTCTAAGGCACAGAAAGTTCCTGAGGAGGAT 1758
QY 1504 GGATCATGCAAGATGGCACACCATGGCCAGGAAAAAATACACAGGGATCATCTCGAATGA 1563
Db 1759 GGATCATGCAAGATGGCACACCATGGCCAGGAAAAAATACACAGGGACCATCTCGAATGA 1818
QY 1564 TTCAGGTTTCTTGGTCAAGTGTGGCTTGTATCTAGAGGTAAATGAGTCCCGCTT 1623
Db 1819 TTCAGGTTTCTTGGTCAAGTGTGGCTTGTATCTAGAGGTAAATGAGTACCGCTT 1878
QY 1624 TAGTTATGTCCTCGTGAAGAGCTCTGGGTCAGACACCAAGAGGCTGGTGCAC 1683
Db 1879 TGGTCTATGTTTCTCGTGAAGAGCTCTGGATTCAGCATCACAAAGCTGGTGCCA 1938
QY 1684 TGAATGCCCTTGTTCGTGTCTCAGCTGTCTTACTTAATGGAACAATCATGTTGAATCTTG 1743
Db 1939 TGAATGCTCTTGTTCGTGTCTCAGCTGTCTTACCAATGGAACAATCATGTTGAATCTTG 1998
QY 1744 ATTGTGATCACTATACATCAACACAGCAAGCTGTCCGAGAGCTATGTCTTCCTAATGG 1803
Db 1999 ATTGTGATCACTATACATCAACACAGTAAGGCTCTCAGGGAAGCTATGTCTTCTTATGG 2058
QY 1804 ATCCAAACCTAGGTCGCGCAAGTCTGTATGTGSCAGTCCCAAAAGGTTGATGGATTG 1863
Db 2059 ACCCTAACCTAGGAAGGAGTGTCTCTAGCTCCAGTTTCCCAGAGATTCGATGSCATTG 2118
QY 1864 ATAGGAATGATCGATATGCAAAACAGGAACACTGTCTTTTGTGATATTAACTTGAGGGGCC 1923
Db 2119 ACAGGAATGATCGATATGCAACACAGGAACACCGTGTCTTCGATATTAACTTGAGGGTC 2178
QY 1924 TTGACGGCATTCAGGACAGTTCATGTTGGAACATGTTGTTGTTTCAACAGAAACAGCTA 1983
Db 2179 TTGATGGCATCAAGGACCAGTTCATGTCGGAACCTGCTGTTTTCACACCGAACAGCTC 2238
QY 1984 TCTATGTTTATGAGCCGCCAAATTAAGGCGAAGAACCGAGGTTTCTTGTGCATCACATATG 2043
Db 2239 TATATGTTTATGAGCCCCCAATTAGCAGAGAGAGGTTGTTTCTTGTCATCACATATG 2298
QY 2044 GGGCAAGAGAAAGCAAGCAAGTCAAGAAAGAGGCTCAGATAGAAAGTCCAAACA 2103
Db 2299 GCGGTAGGAAGAGGCAAGCAAAATCAAGAA---GGGCTCGGACAGAAAGTCCGAGA 2355
QY 2104 AGCATGTGACAGTTCCTGTTCCAGTATTCATCTCGAAGACATAGAGAGGTTGTTGAAG 2163
Db 2356 AGCATGTGACAGTTCCTGTCGACGATTTCAACCTTTGAAGATATAGAGGAGGTTGAAG 2415
QY 2164 GTGCTGGGTTTGATGATGAGAAATCAGTTCATGCTCTCAATGAGCTTAGAGAGAGAT 2223
Db 2416 GCGCTGGATTGACGACGAGAAATCACTTCTTATGCTCAAAATGAGCCTGGAGAGAGAT 2475
QY 2224 TTGGCCAGTCAGCAGCATTTGTTGCTCCACTCTGATGGAATATGTTGGTGTCTCCTCAGT 2283
Db 2476 TTGGCCAGTCCGACGCTTGTGTGCTCCACTCTGATGAGTATGTTGGTGTCTCCTCAGT 2535
QY 2284 CCTCCACTCCAGATCTCTTTTGAAGAGAGCTATCCATGTCATAGTTGTGGCTATGAGG 2343
Db 2536 CCGCAACTCCGAGTCTCTTCTGAAAGAGAGCTATCCATGTTTATAAGCTGTGGCTATGAGG 2595
QY 2344 ACAAGTCTGAATGGGAACTCAGATTTGGTGGATCTGATGATCTGTACAGAGATATTC 2403
Db 2596 ACAAGACTGAATGGGAACTGAGATTCGGGTGGATCTACGGTTCTGTGACAGAGACATTC 2655
QY 2404 TTACTGGATTCAAGATGCACGCAAGAGGCTGGCGTTCACTCTATTTGATGCCAAGCGCC 2463
Db 2656 TCACGGATTCAAGATGCACGCGGAGGCTGGCGGTGATCTACTGATGATGCCAAGCGCC 2715
QY 2464 CAGCTTTCAAGGGATCTGCCGCCCATCAATCTTTTTCAGATCTGTGATCCAGTGTGCGGT 2523
Db 2716 CAGCTTTCAAGGGTCTGCCGCCCATCAATCTTTTTCAGACCGCTCTGAACCCAGGTGCTCCGT 2775
QY 2524 GGGCTCTCGGTTCTGTTGAAATTTCTTTTTCAGCCCGGCATTTGCCCTTATGTTGATGGCTACG 2583

Db 2776 GGGCTTGGTCCGTGGAGATCTCTTACGGGCACTGCCCCCTGTGTACGGGTACG 2835
QY 2584 GAGGGCCCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACTTACCCACTAA 2643
Db 2836 GAGGGCCGTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACTTACCCGCTCA 2895
QY 2644 CTTCTCCGGCTTCTAGTCTATTGTATATGCTGCTATCTGTCTGCTCACTGGAAAGT 2703
Db 2896 CGTCCATCCGGCTTCTCATCTACTGCTATCTGCGGCCATCTGTCTGCTCACCGAAAGT 2955
QY 2704 TCATCATGCGCAGAGATTAGCACTTGGCCAGTATCTGGTTTCAATGCGCTCTTCTTTCAA 2763
Db 2956 TCATCATGTCAGAGATCAGCACTTCCGACAGATCTGGTTTCACTCCCTTTCATCTCGA 3015
QY 2764 TTTTCCCACTGATTCCTTGAGATGAGGTGGAGTGGTGTGTCATGTAGAGTGTGGA 2823
Db 3016 TCTTGGCCACGGGCATCTCGAGATGAGGTGGAGCGGGTGGGCATCGACGAGTGGTGA 3075
QY 2824 GGAATGAAAGTCTCGGTTCATTTGAGGTATCTCTGCACATCTGTTTGGCGTCTTTCAGG 2883
Db 3076 GGAACGAGAGTCTCGGTGATCGGGGCATCTCCGCGACCTCTTTCGCGGTGTTCCAGG 3135
QY 2884 GTCTTCTGAAGTGTCTGCGGTATCGACACCAATCTTCACTGTCACTCAAAAGGCTAATG 2943
Db 3136 GCTGTCTCAAGTGTCTGCGGCATCGACACCAATCTTCACTGTCACTCAAAAGGCTCGG 3195
QY 2944 ACAAGAGCGGACTTTGTGAGCTCTACATGTTCAAGTGGAGAGAGTCTTCACTCCCTC 3003
Db 3196 ACAGAGACGGCACTTTCGCGGAGCTGTACATGTTCAAGTGGAGAGAGTCTTCACTCCCGC 3255
QY 3004 CGAGACCACTTTTGAATTAACATGTTGGTGTCTGTTGTCGACCTCTACGCCATCA 3063
Db 3256 CCACCACTTCATGATCAACCTGGTGGGTCTGTCGCGGCATCTCTCGGCCATCA 3315
QY 3064 ACAGTGTACCAATCATGGGGCGCTCTTTTGGAGGCTCTTCTTGGCTTCTGGGTGA 3123
Db 3316 ACAGCGGATACCACTGTTGGGCGCGCTCTTCGGCAAGCTCTTCTCGCTTCTGGGTGA 3375
QY 3124 TTGTTCACTTATACCACTTCTCAAGGCTTTATGGGAGGCAAAACCGGACACACGACGA 3183
Db 3376 TGTTCACCTGTACCGGCTTCTCAAGGCGCTCATGGGAGGCAAAACCGGACACCGGACCA 3435
QY 3184 TTGTCATGCTGGGCTGCTCCTCGTCTTATCTTCTCTGCTGCTGGGTGCTGTTG 3243
Db 3436 TGTGCTGCTGGGCGATCTCTGCTGGGTCATCTTCTCTGCTGCTGGGTGCTGCTG 3495
QY 3244 ATCCATTCACTACCGCTCTGCTGGGCGCAATATCCAAACCTGTGGCATCAACTGTAGG 3303
Db 3496 ACCCTTCAACACCGGCTCACTGGCCGGATACCCAGAGCTGTGGCATCAACTGTAGG 3555
QY 3304 AAGTGGGAGTTT 3316
Db 3556 GAAGTGGAGGTT 3568

RESULT 7

US-10-160-719-49

; Sequence 49, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822

; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 49
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3449)
US-10-160-719-49
Query Match 70.9%; Score 2572.6; DB 15; Length 3746;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2865; Conservative 0; Mismatches 439; Indels 9; Gaps 2;
QY 4 CGAGGAACCCGCTCCAGCTGTGTCTCGTGGTGGATCGCTCTGCGCGGCGGAGCTGTGCGCATG 63
Db 265 CCAGGCCCCAGGCTCCAGGCCAGCTCCCTCGAGCTTCTCGGCGAGCTGCTTGGCCATGG 324
QY 64 ACGGCGACGGGACGCCCTGAGTCCGGGAGGACACGGGCGCGGAGAGCTGTGCGCATGCT 123
Db 325 AGGGCGACGGGACGCCGCTGAGTCCGGGAGGCGCGTGGCGGACAGGTGTGCGCATGCT 384
QY 124 GCGCCGACGGCTGGGCAACACGTTGGACGGCGACGCTTTCACCGCTCTGCGAGCTGTGCC 183
Db 385 GCGGCGACGGCTGGGCAACACGCGGAGGGGACGCTTTCGCGGCTGTGCGAGCTGTGCG 444
QY 184 GCTTCCGGTCTGCGGCCCTCTGCTAGGACGAGCGCAAGAGGGGACCCAGGCTGTGCC 243
Db 445 GGTTCCTGGTGTGCGGCCCTCTGCTAGGATGACGAGCGCAAGGACGGGACGCGCATGCC 504
QY 244 TCAGTGCAGAGCAAGTACAGCGCCACAGAGGGAGCCGAGGATCCGCGGGAGGAGAG 303
Db 505 CCAGTGCAGAGCAAGTACAGCGCCACAGGGGAGCCGCGGATCCGTTGGGGAGAGAG 564
QY 304 GCGACGACACTGATGCCGATGATGTGTAGTACCTTCACTACCTGTGCTGTGGCACTGAGG 363
Db 565 GAGACGACACTGATGCCGAGTGTGTAGTACCTTCACTTACCTTGTGCTGTGGCACTGAGG 618
QY 364 ACCAGAGCAGAGATGCTGTGAGAGTGGCGAGTGGCGCATGAACACCGGGGCGAGTG 423
Db 619 ACCAGAGCAGAGATGCTGTGAGAGTGGCGAGTGGCGCATGAACAGTGTGGGGGCGAGG 678
QY 424 GCAATGCTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCTCTCCAAGTATGACAGTG 483
Db 679 GGGATGTTGTGCTGCCCAAGTATGACAGTGGCGAGATCGGCTTACCAAGTATGACAGTG 738
QY 484 GAGAGATCCCTAGGGGATAGTCCCTTCACTACCAACAGCCAGATGTTCAGGAGAAATCC 543
Db 739 GCGAGATTCCTCGGGGATACATCCCATCAGTCACTAACAGCCAGATCTCAGGAGAAATCC 798
QY 544 CTGGAGCTTCCCTGTGATTCATCATGATGCTCCCTTACGGGGAAACATCAGCAGACGTGCTC 603
Db 799 CTGGTGTCTTCCCTGACCATCATATGATGTCTCCCACTGGGAAACATTTGGCAAGCGTCTC 858
QY 604 CGTTTCCCTATGTGAATCATTCACCAATCCGTCAGGGAGTGTCTCGGCGATTTGGGA 663
Db 859 CATTTCCCTATGTGAACCATTCGCCAATCCGTCAGGGAGTGTCTCTGTGTAGCATTTGGGA 918
QY 664 ATGTTGCTGGAAAGAGAGATTTGATGGCTGGAAAAATGAAGCAGGAGCAAGGGGTGCGATTC 723
Db 919 ATGTTGCTGGAAAGAGAGAGTGTGATGGTGTGAAAAATGAAGCAGGAGCAAGGGGACGATTC 978
QY 724 CATGACTAATGGGCAAGCATTTGCTCCCTTGAAGTTCGGGCGAGCTACTGACATCGATG 783
Db 979 CCATGACGAATGGCAAGCATTTGCTCCCTCTCAGGGTTCGGGTGTGGTGTGATTTGATG 1038
QY 784 CATCTACTGAATACAAATGGAAGAGCGCTTTACTGAATGATGAACCTCGCCAGCTCTAT 843
Db 1039 CATCACTGATTAACATGGAAGATGCTTTATTTGAACGACGAACCTCGACAGCTCTAT 1098
QY 844 CTAGAAAGTCCCATTTGCTTCTCCAAATAAATCCCTACAGAAATGCTCAATGTTCTGTCG 903

1099 CTAGGAAAGTCCACTTCTCCTCAGGATAAATCCATACAGGATGGTCAATGTGCTGC 1158
904 GGTGTTGTTCTAAGCATCTTCCTGCACACTACCGTCTCAAAATCCTGCGCTAATGCAT 963
1159 GAATGATGTTCTAAGCATCTTCCTGCACACTACCGTATCAAAATCCTGCGCAATGCAT 1218
964 ACCCACTGGGCTTTATCTGTTATATGTAGATTTGGTTTCTTTATCCTCGATACTGG 1023
1219 ACCCATATGGCTTCTATCTGTTATATGTAGATCTGGTTTCTTTCTGCGATATTGG 1278
1024 ATCACTCCGAAAGTGGTTTCCAAATCAACCGGGAGACCTACCTTGATAGACTGGCTTTAA 1083
1279 ATCACTTCCCTAAGTGGTTTCCAAATCAACCGGGAGAGGTACCTTGATAGGCTGGCATTA 1338
1084 GGTATGACCGAAGAGGTGAACCGTCTCAGTTGGCTGCTGTTGACATATATTTGTCAGTACAG 1143
1339 GGTATGACCGGAGAGGTGAGCCATCTCAGTTGGCTGCTGTTGACATATTTGTCAGTACAG 1398
1144 TCGACCCCTTGAAGAGACCACTATCTGTCACCTGCAACACACTGCTGCTATCCATTTCTGCTG 1203
1399 TCGACCCCAATGAGAGAGCTCTCTGTCACCTGCAATACCGTGTCTATCCATCTCTGCTG 1458
1204 TTGATTAATCCCGTGACAAAGGTCTCTTGCTATGTAATCTGATGACGAGGCTTCAATGCTGA 1263
1459 TGGATTACCTGTGATAAGGTCTCTTGCTATGTAATCTGATGAGTGAGTGGATGCTGA 1518
1264 CTTTGTAGCGCATGCTGAGACTTCAGACTTCAGAGTTTGTAGGAATATGGTACATTTGTGAAGA 1323
1519 CATTTGATGCACTAGCTGAGACTTCAGAGTTTGTAGAAAATGGGTACCATTTGTTAAGA 1578
1324 AGTATGACATTTGAACCCAGAGCTCCGAGTTTACTTTTGGCCAGAAAATGATTAACCTGA 1383
1579 AGTACAAATTTGAACCTAGAGCTCTGTAATGGTACTTCTCCAGAAAATGTAATTACTTGA 1638
1384 AAGCAAAAGTCCAGCTTCAATTTGTTAAAGACCGCCGGGCCCATGAAGAGANAATATGAAG 1443
1639 AGGACAAAGTGCACCTTCAATTTGTTAAAGACCGCCGGGCCCATGAAGAGANAATATGAAG 1698
1444 AATTTAAATCAGGANAATGCCCTAGTTTCTAAGGCATTTGAAGTCCCGAGGAAGGAT 1503
1699 AATTTAAAGTTAGGGTAAATGGCTTGTGTTAAGGCACAGAAAGTTCTCGAGGAAGGAT 1758
1504 GGATCATCAAGATGGCACACATCGGCCAGAGAAACAATACAGGAGATCATCTCGAATGA 1563
1759 GGATCATGCAAGTGGCACACATCGGCCAGAGAAACAATACAGGAGCATCTCGAATGA 1818
1564 TTCAGGTTTTCCTTGGTCACAGTGGTGGCTTGATCTGATGAGGTAATGAGCTCCCGCGTT 1623
1819 TTCAGGTTTTCCTTGGTCACAGTGGTGGCTTGATCTGATGAGGCAATGAGCTACCCCGTT 1878
1624 TAGTTTATGTCCTCGTGAAAGGTCCTGGTTCCAGCACACAGAGAGGCTGGTGCCA 1683
1879 TGGTCTATGTTTCTCGTGAAAGGTCCTGGATTCCAGCATCAACAAGAAAGCTGGTGCCA 1938
1684 TGAATGCCCTTGTTCGTCTCAGCTGCTCCTTACTTAATGGAACAATACATGTTGTAATCTTG 1743
1939 TGAATGCTCTTGTTCGTCTCAGCTGTCCTTACATGGAACAATACATGTTGTAATCTTG 1998
1744 ATTGTGATCACTACATCAACAACAGCAGGCTGTCGAGAGAGCTATGTGCTTCTCTAATGG 1803
1999 ATTGTGATCACTACATTAACAACAGTAAGGCTCTCAGGAGAGCTATGTGCTTCTCTAATGG 2058
1804 ATCCAAACCTAGGTCGCGAAGTCTGTTATGTGAGTGTCCCAAAAGGTTTGAATGGGATTG 1863
2059 ACCCTAACCTAGGAGAGGAGTGTCTGCTACGTCGACGTTTCCCGACAGAGATTCGATGGCATTG 2118
1864 ATAGGAATGATCGATATGCAAAACAGGAACAATGCTCTTTTGTATTAATTAACCTTGAAGGCC 1923
2119 ACAGGAATGATCGATATGCCAAACAGGAACAACCGTGTGTTTTCGATTAATTAACCTTGAAGGTC 2178
1924 TTGACGGCAATTAAGGACCACTTTATGTGGAACTGGTTGTTTTCGAAACAGACAGCTA 1983

2179 TTGATGGCATCAAGGACCAAGTTTATGTATGTCGAACTGGCTGTGTTTTCACCGAACAGCTC 2238
1984 TCTATGTTATGAGCCCCCAATTAAGGCGAAGAACCAAGGTTTCTTGGCATCACTATGTG 2043
2239 TATATGTTATGAGCCCCCAATTAAGCAGAAAGAGGTTGTTTCTTGTCTCATCACTATGTG 2298
2044 GGGGCAAGAAAGCAAGCAAGTCAAAAGAAAGAGGCTCAGATAAGAAAAGTTCGAACA 2103
2299 GCGGTAGGAAGAAAGCAAGCAAAATCAAGAA--GGGCTCGGACAAAGAAAGTCCGAGA 2355
2104 AGCATGTGCAAGTCTCTGTTCCAGTATTCAATCTCGAAGACATAGAGAGGGTGTGTTGAAG 2163
2356 AGCATGTGCAAGTCTCTGTTGCCAGTATTCAACCTTGAAGATATAGAGAGGAGTGTGAAG 2415
2164 GTGCTGGGTTTGATGATGAGAAATCAGTCTCTCATGCTCAATATGAGCTTGAAGAAAGAT 2223
2416 GCGCTGGATTGTGACGACGAGAAATCACTTCTTATGCTCTCAATGAGCCTGGGAAAGAT 2475
2224 TTGSCCAGTCAAGCAGCATTTGCTGCCTCCACTCTGATGGAATATGGTGGTGTCTCTCAGT 2283
2476 TTGSCCAGTCCGACGCTGTTGTTGCCCTCACTCTGATGAGATATGGTGGTGTCTCTCAGT 2535
2284 CTTCCACTCCAGAAATCTCTTTTGAAGAAAGCTATCCATGTCATAAGTTGTGGCTATGAGG 2343
2536 CGCAACTCCGAGTCTCTTCTGAAGAAAGCTATCCATGTTTATAGCTGTGGCTATGAGG 2595
2344 ACAAGTCTGAATGGGAAATCAGATTTGGTTGGATCTATGGATCTGTCTCAGAGAAATTC 2403
2596 ACAAGACTGAATGGGAAATCAGATCGGTTGGATCTACGGTCTCTGTGACAGAAACATTC 2655
2404 TTACTGGATTCAAGATGACGACAGAGGCTGGCTTCACTGATGATGATGATGATGATGATGATG 2463
2656 TCACCGGATTCAGAGTGCACGCGGAGGCTGGCGGTGATCTACTGCAATGCCCAAGCGGC 2715
2464 CAGCTTTCAAGGGATCTGCCCCCATCAATCTTTTCAAGATCTGTAACCAAGTGTGCGGT 2523
2716 CAGCTTTCAAGGGTCTGCCCCCATCAATCTTTTGGACCGTCTGACACAGGTTGCTCCGT 2775
2524 GGGCTCTCGGTTCTGTTGAAATCTTTTACCGCGCATTTGCCCTTATGGTATGCTACG 2583
2776 GGGCTCTGGGTCGCGGAGATCTCTTACCGCGCATCTGCCCCCTGTGGTACGCGCTACG 2835
2584 GAGGCGCTCAAGTCTCTGAGAGATTCGCTTACATCAACACACCAATTTACCCATTA 2643
2836 GAGGCGGCTCAAGTCTCTGAGAGATTCGCTTACATCAACACCAATTTACCCATTA 2895
2644 CTTCTCTCCGCTTCTAGTCTATTTGATATTTGCTGCTGCTATCTGCTGCTCACTGGAAGT 2703
2896 GGTCCATCCGCTTCTCATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2955
2704 TCATCATGCGAGATTTAGCAACTTTGGCCAGTATCTGCTTCAATGGCTCTTCTCTTCA 2763
2956 TCATCATTTCCAGAGATCAGCAACTTTGCGCAGCATCTGGTTCATCTCTCTCTCTCTCGA 3015
2764 TTTTTCGCGCTGCTGCTGAGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2823
3016 TCTTTCGCGCAGGCTGCTGAGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3075
2824 GGAATGAACAGTCTCTGGTCAATTTGGAGTATCTCTGACATCTCTGCTGCTGCTGCTGCTGCTGCT 2883
3076 GGAACGAGAGTCTCTGGTGTATCTGGGGCATCTCGCGCATCTCTGCTGCTGCTGCTGCTGCTGCT 3135
2884 GTCTTCTGAAGTGTCTGCGGTATCGACCAACTTCTACTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 2943
3136 GCCTGTCTCAAGTGTCTGCGCGCATCGACCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3195
2944 ACGNAGAGGCTGCTGCTGAGCTTACATGTTCAAGTGGACGACGCTCTCTATCTCTCTCTCTCTCT 3003
3196 ACGAGGACGCTGCTGCGGAGCTGACATGTTTCAAGTGGACGACGCTCTCTGCTGCTGCTGCTGCTGCT 3255
3004 CGACGACATTTTGTATTAACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3063
3256 CCACCACTCTGATCATCAACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3315

; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
US-10-209-059-9

Query Match 70.9%; Score 2572.6; DB 15; Length 3773;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2865; Conservative 0; Mismatches 439; Indels 9; Gaps 2;

QY	4	CGAGGAACCCCGCTCCAGCTCTGCTCGGTGCGGTGGATCGCTCTGCGCGCCATGG	63
DB	282	CGAGGCCAGGCTCCAGGCCAGCTCCCTCGACGTTTCTCGGCGAGCTCGTTGCCATGG	341
QY	64	ACGGGACGCGGACGCCCTGAATCGGGAGGACCGGGGCGGGGAGGTGCGCATCT	123
DB	342	AGGGCGACGCGGACGGCGTGAATCGGGGAGGCGGTGGCGGACAGGTGTGCCAGATCT	401
QY	124	GCGCGACGCGCTGGGACACACCTTGGACGGCGACGCTTTCACCGCTCGACGCTCTGCC	183
DB	402	GCGGCGACGCGCTGGGACACACCGCGGAGGGGACGCTTTCGCGCGCTCGGACGCTCGG	461
QY	184	GCTTCCGCGCTCGCGCCCTGCTACGACGACGAGCGGCAAGGAGGCGACCCAGCGCTGCC	243
DB	462	GGTTTCGGGTGCGCGCCCTGCTACGAGTACGAGCGGAGGACGCGACGCGCGTGCC	521
QY	244	TCGAGTGCAGACCAAGTACAAGGCGCACAGGGAGCCCGAGCGATCCGCGGGAGGAAG	303
DB	522	CCGAGTGCAGACCAAGTACAAGGCGCACAGGGAGCCCGGCGATCCGTTGGGGAGGAAG	581
QY	304	GCGACGACGCTGATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	363
DB	582	GAGACGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	635
QY	364	ACCAGAGCAGAGATGCTGACAGGATGCGAGCTGGCGCATGAACACCGGGGCGAGTG	423
DB	636	ACCAGAGCAGAGATGCTGCGACAGATGCGAGCTGGCGCATGAACGTTGGGGCGAGCG	695
QY	424	GCAATGTTGCCACCCCAAGTATGACAGTGGCGAGATCGGCTCTCCAAATATGACAGTG	483
DB	696	GGGATGTTGGTCCGCCCAAGTATGACAGTGGCGAGATCGGCTTACCAAGTATGACAGTG	755
QY	484	GAGAGATCCCTAGGGGATAGTCCCTTCAGTCCACACGCGAGATGTCAGGAGAAATCC	543
DB	756	GCAGATTCCTCGGGGATACATCCCATCAGTCACTACACGCGAGATCTCAGGAGAAATCC	815
QY	544	CTGGAGCTTCGCTGATCATCACATGATGTCCTTACGGGGAACATCAGCAGAGCTGCTC	603
DB	816	CTGGTCTTCCCTGACCATCATATGATGTCCTTACGGGGAACATCAGCAGAGCTGCTC	875
QY	604	CGTTTCCCTATGGAATCATCAACATCGGTCAAGGAGTTCCTCGGCGAGTATGGGA	663
DB	876	CATTTCCCTATGGAACCATTCGCGCAATCCGTCAAGGAGTTCCTCGGCGAGTATGGGA	935
QY	664	ATGTTGCTCGAAGAGAGAGTTCATGGCTGGAATGAAGCAGGACAAAGGTGCGGATTC	723
DB	936	ATGTTGCTCGAAGAGAGAGTTCATGGCTGGAATGAAGCAGGACAAAGGTGCGGATTC	995
QY	724	CCATGACTAATGGGACAGCATTCCTCCCTCTGAAGTGGGCGAGCTACTGACATCGATG	783
DB	996	CCATGACGAATGGCAGAGCATTCCTCCCTCTGAAGTGGGCGAGCTACTGACATCGATG	1055
QY	784	CATCTACTGATACAAATGGAAGACGCTTACTGAATGATGAACCTCGGCGAGCTCTAT	843
DB	1056	CATCACTGATTAACAATGGAAGATGCTTTATTTGAACGACGAAATTCGACAGCCTCTAT	1115

QY	844	CTAGAAAAGTCCCCATTGCTCTCCAAAATAAATCCCTACAGAAATGGTCATTGTTCTGC	903
DB	1116	CTAGGAAAGTCCCACTTCTCTCCAGGATAAATCCATACAGGATGGTCATTGTGCTGC	1175
QY	904	GGTGGTGTCTTAAGCATCTTCCTGCACTACGGTCTCAAAATCCTGTGCGTAATGCAT	963
DB	1176	GATTGATTGTCTTAAGCATCTTCCTGCACTACGGTATCAAAATCCTGTGCGCAATGCAT	1235
QY	964	ACCCACTGTGGCTTTTATCTGTTATATGAGATTGGTTTGTCTTATCTCTGATACTGG	1023
DB	1236	ACCCATTATGGCTTCTATCTGTTATATGAGATTGGTTTGTCTTATCTCTGATACTGG	1295
QY	1024	ATCAGTTCGCCGAAAGTGGTTTCCAATCAACCGGAGACCTACCTTGATAGACTGGCTTAA	1083
DB	1296	ATCAGTTCGCCGAAAGTGGTTTCCAATCAACCGGAGACCTACCTTGATAGACTGGCTTAA	1355
QY	1084	GGTATGACGAGAGGTGAACCGTCTCAGTTGGTCTGTGACATATTTGTCAGTACAG	1143
DB	1356	GGTATGACGAGAGGTGAACCGTCTCAGTTGGTCTGTGACATATTTGTCAGTACAG	1415
QY	1144	TCGACCCCTTGAAGGAGGCCACCTATCGTCACTGCCAACACTGTGCTATCCATTCTTGTG	1203
DB	1416	TCGACCCCAATGAAGGAGCCCTCTCTTGTCTCAGTCCCAATACCGTGTCTATCCATTCTTGTG	1475
QY	1204	TTGATTATCCCGTGGCAAGGCTCTTGTCTATGATATCTGATGACGAGCTTCAATGCTGA	1263
DB	1476	TGGATTACCTCTGGGATAAGGCTCTCTGCTATGATATCTGATGAGGCTGCGATGCTGA	1535
QY	1264	CTTTTGACGATTTGGCTGAGACTTCAAGTTCCTAGGAAATGGTACCAATTTGTGAAGA	1323
DB	1536	CATTGATGACCTAGCTGAGACTTCAAGTTCCTAGGAAATGGTACCAATTTGTGAAGA	1595
QY	1324	AGTATGACATTTGAACCCAGAGCTCCCGAGTTTCTTCTTCCAGAAAATGATTACCTGA	1383
DB	1596	AGTATGACATTTGAACCCAGAGCTCCCGAGTTTCTTCTTCCAGAAAATGATTACCTGA	1655
QY	1384	AAGACAAAGTCCAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1443
DB	1656	AGGACAAAGTCCAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1715
QY	1444	AATTTAAATCAGGATTAATCCCTAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1503
DB	1716	AATTTAAATCAGGATTAATCCCTAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1775
QY	1504	GGATCATGCAAGATGCGACACCATGCGCGGCAACCAATACAGGAGTATCTCTGGAATGA	1563
DB	1776	GGATCATGCAAGATGCGACACCATGCGCGGCAACCAATACAGGAGTATCTCTGGAATGA	1835
QY	1564	TTGAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1623
DB	1836	TTGAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1895
QY	1624	TAGTTTATGCTCTCGTGAAGAGCTCTCGGTTTCCAGACCAACCAAGAGCTGGTGCCA	1683
DB	1896	TAGTTTATGCTCTCGTGAAGAGCTCTCGGTTTCCAGACCAACCAAGAGCTGGTGCCA	1955
QY	1684	TGAATGCCCTTGTCTGCTGCTCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1743
DB	1956	TGAATGCCCTTGTCTGCTGCTCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2015
QY	1744	ATTGTGATCATCATCAACACAGCAGCTGTCCGAGAGCTGTCCGAGAGCTATGCTTCTTATGG	1803
DB	2016	ATTGTGATCATCATCATCAACACAGCAGCTGTCCGAGAGCTGTCCGAGAGCTATGCTTCTTATGG	2075
QY	1804	ATCCAACTAGTTCGCGAGTCTGTTATGTCAGTTTCCCAAGAGTTTGTATGGGATG	1863
DB	2076	ATCCAACTAGTTCGCGAGTCTGTTATGTCAGTTTCCCAAGAGTTTGTATGGGATG	2135
QY	1864	ATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1923
DB	2136	ATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2195
QY	1924	TTGACGGCATTCAGGACCGAGTTTATGTTGGAACTGGTTGTTTCTTCTTCTTCTTCTTCTT	1983

Db 2196 TTGATGGCATCAAGGACAGTTTATGTCGGAACTGGCTGTGTTTCAACCGAAGCTC 2255
QY 1984 TCTATGGTTATAGAGCCCAATTAAGGCGAAGAGGTTCTTGGGCAATCACTATGTG 2043
Db 2256 TATAATGGTTATAGAGCCCAATTAAGGCGAAGAGGTTGGTTCTTGTCTATCACTATGTG 2315
QY 2044 GGGCGAAGAAGCAAGCAAGTCAAGAAAGAGGCTCAGATAGAAAAGTCAACA 2103
Db 2316 GCGGTAGAGAGAGCAAGCAATCAAGAA--GGGCTCGGACAGAAGAAGTGGCAGA 2372
QY 2104 AGCATGTGACAGTTCTGTTCAGTATTCATATCTCGAAGACATAGAGGAGGTTGTGAAG 2163
Db 2373 AGCATGTGACAGTTCTGTGCCAGTATTCACCTTGAAGATATAGAGGAGGTTGAAG 2432
QY 2164 GTGCTGGGTTTGAATGAGAAATCAGTTCTCATGTCTCAATAGAGCTTGAAGAAGAT 2223
Db 2433 GCGCTGGAATTTGACACGAGAAATCACTTCTATGTCTCAATGAGCCTGGAGAAGAT 2492
QY 2224 TTGGCAGTCAAGCAGATTTGTCCTCACTCTCATGGAATATGTTGTTCTCAGT 2283
Db 2493 TTGGCAGTCCGACGTTTGTGCTCACTCTCATGAGGATGTTGGTGTTCCTCAGT 2552
QY 2284 CCTCACTCCAGATCTCTTTGAAGAAGCTATCCATGTCTATAAGTTGTGGCTATGAG 2343
Db 2553 CCGCAACTCCGAGTCTCTCTGAAGAAGCTATCCATGTTATAAGCTGTGCTATGAG 2612
QY 2344 ACAAGTCTGAATGGGAATGAGATGGTTGATCTATGGATCTGTCTCAGAGAATATTC 2403
Db 2613 ACAAGACTGAATGGGAATGAGATCGGTTGATCTACGGTTCTGTGACAGAAGACATTC 2672
QY 2404 TTACTGGATTCAGATGACGACGAGGCTGCGTTTCACTATGTCATGCGCCCAAGCGCC 2463
Db 2673 TCACCGGATTCAGATGACGCGGAGGCTGCGGTTCGATCTACTGCAATGCCCAAGCGC 2732
QY 2464 CAGCTTTCAAGGGATCTGCCCCCATCAATCTTTAGATCTGTGCAACCAAGTGTGCGGT 2523
Db 2733 CAGCTTTCAAGGGTCTGCCCCCATCANCTTTGCGACGGTCTGAAACAGTGTCCGGT 2792
QY 2524 GGGCTCTGGTCTGTTGAAATCTTTTACGCGGCATGCCCCCTATGTTGTTAGCTACG 2583
Db 2793 GGGCTCTGGTCTGGTGGAGATCTCTTTCAGCGGCACCTGCCCCCTGTGTACGCTACG 2852
QY 2584 GAGGCGGCTCAAGTTCCTGGAGATTCGCTTACATCAACACACCACTTTACCCACTAA 2643
Db 2853 GAGGCGGCTCAAGTTCCTGGAGATTCGCGTACATCAACACCACTATCCCGCTCA 2912
QY 2644 CCTCTCTCCGCTTCTAGTCTATTTGATATTTAGTCTGCTGCTGCTCACTGGAAGT 2703
Db 2913 CGTCCATCCGCTTCTCATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2972
QY 2704 TCATATGCCAGAGATTAGCAACTTGGCCAGTATCTGTTCAATGCGGCTTCTTCTTTCAA 2763
Db 2973 TCATATTCAGAGATCAGCAACTTGGCCAGATCTGTTTCAATCTCTCTCTCTCTCTCGA 3032
QY 2764 TTTTCCGCACTGGTATCCTTCAGATGAGTGGAGTGGTGGATGACAGAGTGGTGA 2823
Db 3033 TCTTTCGCAACGGGATCTCTGGAGATGAGTGGAGGCGGTGGGATCAGCAGTGGTGA 3092
QY 2824 GGAATGAACAGTCTTGGGTCAATGGAGTATCTCTGCAATCTGTTTGGCCGCTTCTTCAAG 2883
Db 3093 GGAACGAGCAGTCTCTGGTATCGGGGCACTCTCGGCGCACTCTTTCGCGGTGTTCCAG 3152
QY 2884 GTCTTCTGAAGTCTTCCCGGTATCGACACAACTTCACTGTCACTCTCAAGGCTAATG 2943
Db 3153 GCCTGCTCAAGTGTGCGCGGCACTCGACACAACTTCACTGTCACTCTCAAGGCTCTCG 3212
QY 2944 ACGAAGAGGAGCTTTGCTGAGCTCTACATGTTCAAGTGAACAGCGTCTCTATCTCTC 3003
Db 3213 ACGAGGACGGGACTTTCGCGAGCTGTACATGTTCAAGTGAACAGCGTCTCTATCTCCG 3272
QY 3004 CGACGACCATTTTATCATTAACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3063

Db 3273 CCACCACCATCTGATCATCAACCTGTCGGCGTCTGTCGCCGCGATCTCTCTACGCCATCA 3332
QY 3064 ACAGTGGTTACCAATCATGGGGCGGCTCTTTTGGGAAAGCTCTTTTGGGCTTCTTGGGTGA 3123
Db 3333 ACAGCGGATACAGTCTGTTGGGCGCGCTCTTTCGGCAAGCTCTTCTTGGCTTGGGTGA 3392
QY 3124 TTGTTCACTTATACCATCTCTCAAGGGTCTTATGGGCGAGGCAAAACCGCACACCGACA 3183
Db 3393 TCGTCCACCTGTACCGCTTCTCAAGGGCTCTATGGGAGGAGAACCGCACCCCGATCA 3452
QY 3184 TTGTCATGCTGCGGCTGTCTCCCTGCTCTATCTTCTCTTCTTCTTCTTCTTCTTCTTCT 3243
Db 3453 TCGTCTGCTCTGGGCAATCTGCTGGGCTCATCTTCTCTTCTTCTTCTTCTTCTTCTTCT 3512
QY 3244 ATCCATTCACCTACCGTCTGCTGCGCCCAAAATATCAAACTGTGGCATCACTGCTAGG 3303
Db 3513 ACCCTTCACCAACCGGCTCACTGCGCCCGATACCCAGACGTTGTGGCATCACTGCTAGG 3572
QY 3304 AAAGTGGGAGTTT 3316
Db 3573 GAAAGTGAAGGTT 3585

RESULT 10
US-10-160-719-5
; Sequence 5, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helencjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (338)...(3568)
US-10-160-719-5

Query Match
Best Local Similarity 70.9%; Score 2572.6; DB 15; Length 3773;
Matches 2865; Conservative 0; Mismatches 439; Indels 9; Gaps 2;
QY 4 CGAGGAACCCCGCTCCAGCTCTGTGTCGGTGGGTTGATCTGTCGCCGCGCATGG 63
Db 282 CCAGGCCCCAGGCTCCAGGCGAGCTCCCTCGACGCTTCTCGGCGAGCTCGTTGCCATGG 341
QY 64 ACGCGGACGGGAGCCCTGAGTCCGGAGGACACGGGGCCGGGACGTTGTGCCAGATCT 123
Db 342 AGGCGGACGGGAGCGGCGTGAAGTCCGGGAGGCGCGTGGCGGACAGGTGTGCCAGATCT 401
QY 124 GCGCGGACGGCTGGGGACCAACGCTTGGACCGGCGACGCTTTCACCGCTCGGAGCTGCTGCC 183
Db 402 GCGCGGACGGCTGGGGACCAACGCGGAGGGGAGCGTCTTCGCGCGCTCGAGCTCTGCG 461
QY 184 GTTTCGCGTCTGCGCGCTTGTCTACGAGCAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 243
Db 462 GTTTCGCGTCTGCGCGCTTGTCTACGAGTACGAGCGCAAGGAGCGGACGCGACGCGCTGCC 521
QY 244 TCCAGTCAAGACCAAGTACAGCGGCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 303


```
QY 2464 CAGCTTTCAAGGATCTGCCCCATCAATCTTTTCAGATCGTGTGAACCAAGTGTGCGGT 2523
Db |||||
QY 2733 CAGCTTTCAAGGGTCTGCCCCATCAATCTTTTCGACCGTGTGAACCAAGTGTGCGGT 2792
Db |||||
QY 2524 GGGCTCTGGTCTCTGTGAAATCTTTTCAGCGGCAATGCCCTTTATGTTGCTAGC 2583
Db |||||
QY 2793 GGGCTCTGGTCTGGTGGAGATCTCTTCAGCGGCACTGCCCTGTGTGCTAGGCTAGC 2852
QY 2584 GAGGGCGCTCAAGTCTCTGGAGAGATTCGCTTACATCAACACCACTTTACCACTAA 2643
Db |||||
QY 2853 GAGGGCGCTCAAGTCTCTGGAGAGATTCGCTTACATCAACACCACTTTACCACTAA 2912
QY 2644 CTTCTCTCCGCTCTTAGTCTTATTTGATATTCCTGCTATCTGCTGCTCACTGAAAGT 2703
Db |||||
QY 2913 CGTCCATCCGCTTCTCATCTACTGATCTCTGCGCCGCACTGCTGCTCAACCGAAAGT 2972
QY 2704 TCATCATGCGAGATTTAGCAACTTTGGCCAGTATCTGTTGCTATCTGCTCTCTTTCAA 2763
Db |||||
QY 2973 TCATCATTTCCAGAGATCAGCAACTTTGCGCAGCATCTGGTTGATCTCTCTCTTCACTCGA 3032
QY 2764 TTTTCGCCACTGCTATCTCTGAGATGAGTGTGAGTGTGTTGGCATTTGAGAGTGTGGA 2823
Db |||||
QY 3033 TCTTCGCCACGGGCACTCTGAGATGAGTGTGAGCGGGTGGGCATCGACGAGTGTGGA 3092
QY 2824 GGAATGAACAGTCTCTGGGTCTATTTGAGGTATCTCTGCAATCTGTTTGGCGTCTTTGAG 2883
Db |||||
QY 3093 GGAACGAGAGTCTGAGGTATCTGCGGCACTCTGCGGCACTCTTTCGCGGTGTTCCAGG 3152
QY 2884 GTCTTCTGAAGGTCTTCCCGTATCGACACCAACTTCACTGCTCACTCAAAAGGCTAATG 2943
Db |||||
QY 3153 GCCTGCTCAAGGTCTGCGCGCATCGACACCACTTCAACGCTCACTTCCAGGCTCGG 3212
QY 2944 ACGAAGAGGCACTTTCTGAGCTTACATGTTCAAGTGTGAGCGCTTCTCATCTCCTC 3003
Db |||||
QY 3213 ACGAGGACGGGCACTTCTCGGAGCTGTATCATGTTTCAAGTGTGACGAGCTCTCTGATCCCGC 3272
QY 3004 CGACGACCAATTTGATCAATTAACATGTTGTTGCTGTTGCTGCGCACTCTTACGCGCATCA 3063
Db |||||
QY 3273 CCACCACTCTGATCATCACTGCTGCGGCTGCTGCGGCACTCTTCTTCTGCGGTCTG 3332
QY 3064 ACAGTGTTCACATCAATCAATGAGGCGCTCTTTTGGGAAGCTCTTCTTCTGCGGTCTG 3123
Db |||||
QY 3333 ACAGCGGATACAGTCTGAGGCGCGCTCTTTCGGAAGCTCTTCTTCTGCGGTCTG 3392
QY 3124 TTGTTCACTTATACCACTTCTCAAGGCTTCTTATGGGCAAGCAAAACCGCAACCGAGA 3183
Db |||||
QY 3393 TCGTCCACTCTACCGCTCTCAAGGCGCTCATGGGCAAGCAAGCAAGCAAGCAAGCA 3452
QY 3184 TTGTCATGCTGCGGCTCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3243
Db |||||
QY 3453 TCGTCTGCTCTGCGGCACTCTGCTGCGGTCTCACTCTCTCTCTCTCTCTCTCTCTCT 3512
QY 3244 ATCCATTCATACCGCTCTGCGGCAAGCAATCAAACTGTGGCACTCACTGCTAGG 3303
Db |||||
QY 3513 ACCCTTTCAACACCGGCTCACTGCGGCGGATACCAAGCTGTGGCACTCACTGCTAGG 3572
QY 3304 AAGTGGGAGTTT 3316
Db |||||
QY 3573 GAAGTGGAAAGTTT 3585
```

RESULT 11

```
US-10-627-132-13
; Sequence 13, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
```

```
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 3704
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-13
```

```
Query Match 64.2%; Score 2326.2; DB 13; Length 3704;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 2687; Conservative 0; Mismatches 543; Indels 9; Gaps 3;

QY 66 GCGCAGCGGAGCGCCCTGAAAGTCCGGGAGGACACGGGCGCGGAGCGTGTGCCAGATCTGC 125
Db |||||
QY 275 GACGCGCGGAGCGCCACCAATTCGGGGAAGCATGTGGCGGGCAGGTGTGCCAGATCTGC 334
Db |||||
QY 126 GCCGAGCGCTTGGGACCAACAGTTTGAACGCGGACGCTTTTCAACGCTGTGAGCTGTGCCG 185
Db |||||
QY 335 GCGCAGCGGCGTGGGACCGCGCGGACCGGACCTCTTTCACCGCTGTGAGCTGTGCCG 394
QY 186 TTCCCGGCTCCGCGCCCTGTACGACGACGAGCGGAGCGGATCCGCGGGGAGGAAGGC 245
Db |||||
QY 395 TTCCCGGCTGGCGCCCATGCTACGAGTACGAGCGCAAGGAGCGGACCGGCGGTGCCG 454
QY 246 CAGTCAAGACCAAGTACAAGCGCCACAGAGGAGCGGATCCGCGGGGAGGAAGGC 305
Db |||||
QY 455 CAGTCAAGACCAAGTACAAGCGCCACAGAGCGGATCCGCGGGGAGGAAGGC 514
QY 306 GACGACCTGATGCGGATGATGAGTGTGATGCTTCACTACCTGTGATCTGTGGCACTGAGGAC 365
Db |||||
QY 515 GAGGATGTGGATGTGAGCGATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 574
QY 366 CAGAAGCAGAAAGATTGCTGACAGGATGCGGAGCTGGCGCATGACACCGGGGAGTGGC 425
Db |||||
QY 575 CAGAAGCAGAAAGATTGCTGAGAGATGCTCACTTGGCGGACAACTCAGTGGGAGTG-- 632
QY 426 AATGTGTGCCACCCCAAGTATGACAGTGGCGAGATCGGCGCTCTTCAAGTATGACAGTGG 485
Db |||||
QY 633 -ATATGTGCGCTGCTAAGTATGACAGCGGTGAATTTGGGCATGGGAAGTATGACAGTGGT 691
QY 486 GAGATCCCTAGGGATAGTCCCTTTCAGTTCACCAACAGCCAGATGTTCAGGAGAAATCCCT 545
Db |||||
QY 692 GAGATCCCTTCTGGATATATCCCGTCACTAACTCATAGCCAGATCTTCAGGAGAGATTCT 751
QY 546 GGAGCTTCGCTGATCATCATCATGATGTCCCTACGGGAGACATCAGCAGAGTGTCTCCG 605
Db |||||
QY 752 GGAGCTTCCCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 811
QY 606 TTTCCCTATGTGAATCAATTCACCAATTCCTCAAGGAGGATTTCTCCGGCAGTATTTGGGA 665
Db |||||
QY 812 TTTCCCTATGTGAATCAATTCCTCAAGGAGGATTTCTCCGGCAGTATTTGGGAAT 871
QY 666 GTTCCCTGGAAAGAGAGAGTGTGCTGGAATGAAGCAGGACAAAGGTTGCAATTTCC 725
Db |||||
QY 872 GTTGCATGGAAAGAGAGGTTGGATGGATGGAAATGA--AGGATAAAGGTGCAATTTCT 928
QY 726 ATGACTAATGGCAACAGTGTCTCTCTGAGGTTGGGCGAGCTACTGACATCGATGCA 785
Db |||||
QY 929 ATGACCAATGGCAACAGTGTCTCTCTGAGGTTGGGCGAGTGTGATGATGATGATGAT 988
QY 786 TCTACTGAATACACATGGAAGAGCGCTTTTACTGAATGATGAATCTGCCAGCTCTTCT 845
Db |||||
QY 989 TCTACTGATTAACATGGAAGATGCTTTTACTGAATGATGAATCTGCCAGCTCTTCT 1048
```


1709	Db	 ATCATG CAGATGGTACACCTTGGCTGGGAACAATACTAGGACCAATCTCGAATGATT	1768
1566	Qy	 CAGGTTTTCTTGGTACAGTGGTGGCCCTGATACAGGTAATCAGCTCCCCGTTTAA	1625
1769	Db	 CAGGTTTTCTGGGTCAAGTGGAGGGCTTGAAGTTGAAGGCATGAATCTCTCGTTTG	1828
1626	Qy	 GTTTATGTGTCGTGAAAACGCTCTGGGTTCCAGCACACAAGAGGCTGGTGCATG	1685
1829	Db	 GTTTATGTGTCGTGAAAACGCTCTGGGATCCACATCACAAGAGGCTGGTGCATG	1888
1686	Qy	 AATGCCCTTGTTCGTGTCCTCAGCTGCTCTAATAATGGAACAATACATGTTGAATCTTGAT	1745
1889	Db	 AATGCACTTGTTCGTGATCAGCTGTCTTACTAATGGGCAATACATGTTGAATCTTGAT	1948
1746	Qy	 TGTGATCACTACATCAACAAGAGGCTGTCGAGAGGCTATGCTCTCTTAATGGAT	1805
1949	Db	 TGTGACCACTACATCAATAATAGCAAGGCTCTTCGAAAGCTATGCTCTCTTATGGAC	2008
1806	Qy	 CCAAACCTTAGTCGCGCAAGTCTGTTATGTGCAAGTCCCAAAAGGTTGATGGGATGAT	1865
2009	Db	 CCAAACCTAGGAAGGATGTCGTTATGTCCATTTCTCAGAGGTTGATGGTATGAT	2068
1866	Qy	 AGGAATGATCGATATGCAAAACAGAAACACTGTCTTTTTGATATTAACCTTGAGGGGCTT	1925
2069	Db	 AGGAATGACCGATATGCAAAACAGAAACACTGTGTTTTTCGATATTAACCTTGAGAGTCTT	2128
1926	Qy	 GACGGCATTAAGGACCAAGTTTATGTGGGAACGTGGTTGTCTTTCACAGAAACAGCTATC	1985
2129	Db	 GACGGCATTAAGGGCCAGTTTATGTGGGAACGTGGTTGTCTTAAACAAACGGGCTTA	2188
1986	Qy	 TATGTTATGAGCCCCCAATTAAGCGAAAGAGCAGGTTCTTGGCATCACTATGTGGG	2045
2189	Db	 TATGTTATGAGCCTCCAGTCAAGAAAAGAGCAGGCTTCTCTTCGCTTTGTGGG	2248
2046	Qy	 GGCAAGAGAAAGGCAAGCAAGTCAAGAAAAGGAGCTCAGATAAGAAAAAGTCCAAACAAG	2105
2249	Db	 GGAGGAAAAGACGTCAAAATCTAAGAA--GAGCTCGGAAAAAGAAAGTCACATAGA	2305
2106	Qy	 CATGTGACAGTTCTGTTCCAGTATTCATCTCGAAGACATAGAGGAGGTTGTTGAAGT	2165
2306	Db	 CACGCAGACAGTTCTGTACCAAGTATTTAATCTCGAAGATATAGAGGAAGGATGTAAGGT	2365
2166	Qy	 GCTGGTTTTGATGATGAGAAATCAGTTCTCATGTCTCAAAATGAGCTTTAGAGAAGAGATTT	2225
2366	Db	 TCTCAGTTTGATGATGAGAAATCGCTGATTAAGTCTCAATGAGCTTTGGAGAAGAGATTT	2425
2226	Qy	 GGCCAGTCAGCAGCATTTGTTTGCCTCCACTCTGATGGAATATGTTGGTTCTCTCAGTCC	2285
2426	Db	 GGCCAGTCAGTGTGTTTGTAGCCCTCTACTCTGATGGAATATGTTGGTTCTCCACAATCT	2485
2286	Qy	 TCCACTCAGAAATCTCTTTTGAAGAAGCTATCATGTCTAAGTTGTGGCTATAGGAC	2345
2486	Db	 GCAATCCAGAGTCTCTTCTGAAAAGAGCTATTCATGTCTCAGCTGTGGCTATAGGAC	2545
2346	Qy	 AAGTCTGAAATGGGAACGTAGATGGTTGGATCTATGGATCTGTCAAGAAAGATATTTCTT	2405
2546	Db	 AAACTGACTGGGAACGTAGATGGTGGATCTATGGTTCTGTTACAGAAGACATTTCTC	2605
2406	Qy	 ACTGGAITCAAGATCAAGCAGAGGCTGGGTTTCAGTCTATTGCATGCCCAAGCGCCA	2465
2606	Db	 ACCGAATCAAGATGATGTCGAGGCTGGCGATCAATCTACTCATGCTGCCTTAAGCGACCA	2665
2466	Qy	 GCTTTCAAGGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTCTCGGTTGG	2525
2666	Db	 GCTTTCAAGGGATCTGCTCTCAATCAACCTTTTCGGATCGTTTGAATCAAGTCTTCGTTGG	2725
2526	Qy	 GCTCTCGGTTCTGTTGAAATCTTTTCAGCGGCAATTCGCCCTTATGTTATGGCTACGGA	2585
2726	Db	 GCTCTTGGTTCCATTGAAATCTTTTCAGCAGGCAATGTCCCATATGTTATGGCTATGGA	2785
2586	Qy	 GGCGCCCTCAAGTCTCTGGAGAGATTGCTTATACATCAACACCAATTTACCCATAACC	2645

Db	2786	GGCCGGCTTAAATTCCTGGAGAGATTGCTTATATCAACAACAATTTATCCACTCACA	2844
Qy	2646	TCTCTCCGCTTCTAGTCTTATTGTATATGCCGTCTATGTCTGCTCATCTGGAAAGTTC	2705
Db	2846	TCAATCCGGCTCCTCTGTACTGCAATATGCCAGCAGTTGTCTTCTCACTGGGAGTTC	2905
Qy	2706	ATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTCAATGGGCTCTCTCTTCAATT	2765
Db	2906	ATCATCCCAAGATTAGTAACCTAGAGAGTGTGGTGTATATCGCTCTTATCTCAATC	2965
Qy	2766	TTGCCCACTGGTATCCTTTGAGATGAGGTGGAGTGGTCTTGGCATTCAGCAGTGGTGGAGG	2825
Db	2966	TTTGCCACTGGTATCCTTTGAGATGAGGTGGAGTGGTGTGGCATTGATGAATGGTGGAGG	3025
Qy	2826	AATGAAAGTTCTGGGTCAATGGAGGTATCTCTGCACATCTGTTTGGCGTCTTTTCAGGTT	2885
Db	3026	AACGAGCAGTTCTGGGTCAATGGTGGTATTTCTGCGCAITTTATTTGGCGTCTTCCAGGTT	3085
Qy	2886	CTTCTGAAGTGTCTGCGGTATCGACACAACTTTCACCTCTCACTCAAAGGCTAATGAC	2945
Db	3086	CTCTGAAGGTGTCTGCTGGTATCGACACAGCTTCACTGTCACTTAAAGGCCACTGAC	3145
Qy	2946	GAAGAAGCGCACTTTGCTGAGCTCTACATGTTCAAAGTGGACGAGCTTCTCATCTCCCTCG	3005
Db	3146	GAAGAAGTGATTTTTCGAGCTCTACATGTTCAAAGTGGACACGCTTCTGTATCCCACCA	3205
Qy	3006	ACGACCAATTTTGATCAATTAATGGTTGGTGTGTTGCTCGCCACTCTACGCCATCAAC	3065
Db	3206	ACCACTATTTTGTATCATCAACTGGTCTGGGTGGTGTGCTCGCTCGCAATTCCTACGCAATCAAT	3265
Qy	3066	AGTGGTTACCAATCATGGGGCCGCTCTTTGGGAAGCTCTTCTTGGCTTCTGGGTGATT	3125
Db	3266	AGCGGTTACAGTCAATGGGACCTTTTTCGGGAAGCTCTTCTTGGCTTCTGGGTGATT	3325
Qy	3126	GTTTCACTTATACCAATTCCTCAAGGTTCTTATGGGCAAGGCAAAACCGCACACCGAGATT	3185
Db	3326	GTCCACCTGTACCCCTTCTCAAGGGCTCATGGGGAAGCAGAACCGCACCGCCACCAATT	3385
Qy	3186	GTATCTGTCTGGGTGTCCTCTCGCTCTATCTTCTCTTGGCTGTGGTTCGTTGAT	3245
Db	3386	GTCTGTGTCTGGGTATCTCTCTTCTGCGTCGATCTTTTCCCTGATGTGGTTCGTATCGAT	3445
Qy	3246	CCATTCACTACCGCTCTCGCTGGGCCAAATATCCAAACCTGTGGCATCAACTCTAGGA	3304
Db	3446	CTATTCACACCGGGTCACTGGCCCTCATATCTCGGAATCTGCAATCAACTCTAGGA	3504

```

RESULT 13
US-10-160-719-57
; Sequence 57, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 57
; LENGTH: 3704
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS

```

i LOCATION: (272)...(3499)
US-10-160-719-57

Query Match 64.2%; Score 2326.2; DB 15; Length 3704;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 2687; Conservative 0; Mismatches 543; Indels 9; Gaps 3;

QY	66	GGCGACGGGACGCCCTGAAGTCCGGGAGGACCGGGCCGGGAGCGTGTGCCAGATCGC	125
DB	275	GACGGCGGAGCGCCACCAATTCGGGGAGCATGTGGCCGGGACGGTGTGCCAGATCGC	334
QY	126	GCCGACGGCCTGGGACCAACAGTGTGACCGCGACGCTTTCACCGCCTGGGACGCTCTGCCGC	185
DB	335	GCGGACGGCGTGGGACCGCGCGGACGGGACCTTTCACCGCCTGGGACGCTCTGCCGC	394
QY	186	TTCCCGGTCTGCGGCCCTGCTACGAGCAGCGGAGGAGGGACCCAGGCGCTGCCTC	245
DB	395	TTCCCGGTCTGCGGCCCTGCTACGAGTACGAGCGCAAGGACCGGACCGCGCTGCCTC	454
QY	246	CAGTCAAGACCAAGTCAAGCGCCACAGAGGAGCGCCAGCGATCCGGGGGAGGAGGC	305
DB	455	CAGTCAAGACCAAGTCAAGCGCCACAGGAGGAGCGCCACCGATACACGGTGAGGAAAT	514
QY	306	GACGACCTGATGCGGATGATGTAGTGACTTCAACTACCTCTGCACTGGCACTGAGGAC	365
DB	515	GAGGATGTGGATGCTGACGATGTGAGTGACTACAACTACCAAGCATCTGGCAACCAAGAT	574
QY	366	CAGAGCAGAGATGTGTGACAGGATGCGCAGCTGGCGATGAACACCGGGGGAGTGGC	425
DB	575	CAGAGCAGAGATGTGTGAGAGATGTCTCACTTGGCGGACAACTCACTGGCGGAGTG--	632
QY	426	AATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTCCCAAGTATGACAGTGA	485
DB	633	-AATGTTGGCCCTGCTAAGTATGACAGCGTGAAATTTGGCATGGGAAGTATGACAGTGGT	691
QY	486	GAGATCCCTAGGGGATACGTCCTTTCAGTCAACCAAGCAGAGATGTGAGGAAATCCCT	545
DB	692	GAGATCCCTGCTGGATATATCCCGTCACTAACTCATAGCCAGATCTCAGGAGAGATTCCT	751
QY	546	GGAGTTCGCTGATCATCATGATGTCCCTACGAGGAGTCCCTACGAGGAGATCATGAGCAGT	605
DB	752	GGAGTTCGCTGATCATCATGATGTCTCTGTGGGAAACATTTGGCAGCGGTGGACATCAA	811
QY	606	TTTCCCTATGTGAATCATTTACCAATCCGTCAAGGGAGTCTCCGGCAGTATTTGGGAT	665
DB	812	TTTCCCTATGTAAATCATTTCCAAACCCATCGAGGAGTCTTCGGGTAGCCCTTGGCAAT	871
QY	666	GTTGCTGGAAAGAGAGATGTGATGGCTGGAAATGAAGCAGGACAAAGGGTGGCAATTCCT	725
DB	872	GTTGCTGGAAAGAGAGGTTGGATGGATGGAAATGA--AGGATAAAGGTGCAATTCCT	928
QY	726	ATGACTAATGGGACAGGATTTCTCCCTCTGAGGTGGGAGGAGTCTGACATGATGCA	785
DB	929	ATGACCAATGGAACAAGCATTTGCTCCATCAGAAGGGCGTGGAGTTGCTGATATTTGATGCT	988
QY	786	TCTACTGAATCAACATGGAAGACGCTTTACTGATGATGAACCTCGCCAGCCTCTACT	845
DB	989	TCTACTGAATTAACATGGAAGATGCTTTACTGATGATGAACCTCGCCAGCCTCTACT	1048
QY	846	AGAAAGTCCCATTTGCTTCCCAAAATAATCCCTACAGAAATGGTCAATTTCTGGCG	905
DB	1049	AGAAAGTGGCAATTTCTTCCATCCAGAAATAATCCGTACAGAAATGGTCAATTTGCTACGT	1108
QY	906	TTGGTTGTTCTAAGCATCTTCTGCACTACCGTCTCAAAATCTCTGGCGTAATGCAATAC	965
DB	1109	TTGGCTGTCTTATGATGATTTCTTGGGCTACCGGTATCACATCTCTGTGAAACAATGCAAT	1168
QY	966	CCACTGTGGCTTTTATCTGTTATATGTGAGATTTGGTTTGTCTTATCTCTGATCTGGAT	1025
DB	1169	CCACTGTGGCTTTTATCTGTTATATGTGAGATTTGGTTTGTCTTATCTCTGATCTGGAT	1228
QY	1026	CAGTTCGCGAAGTGGTTTCCCAATCAACCGGGAGACCTACCTTTGATGAGACTGGCTTTAAGG	1085

DB	1229	CAGTTCGCGAAGTGGTCCCCCAATCAACCGTGAAACATACCTTGTAGACTGGCTTTAAGG	1288
QY	1086	TATGACCGAGAAAGTGAACCGTCTCAGTTGGCTGTGTGACATATTTGTTCAGTACAGTC	1145
DB	1289	TATGACCGAGAAAGTGAACCAATCTCAATAGCTCCTGTGTGATATTTTGTTCAGTACTGTG	1348
QY	1146	GACCCCTTGAAGGAGCCACCTATCGTCACTGCGCAACACATGCTATCCATTTCTGCTGT	1205
DB	1349	GATCCAAATGAAGAGGCCCTCTCTGTCACTGCAAAATACGTGTCTTTCCATCCTTCTGTGC	1408
QY	1206	GATTATCCCGTGGACAAAGTCTCTTGTATGTATCTGATGACGAGCTTCAATGCTGACT	1265
DB	1409	GATTATCCCGTGTGACAAAGTATCTTGTATGTTCGATGATGAGAGCTGCTATGCTGACT	1468
QY	1266	TTTGAACGATTTGGGTGACATTTCAAGATTTGCTAGGAAATGGGTACCAATTTGTGAAGAG	1325
DB	1469	TTTGAATGCTCTCTCTGAAACTTCAGAGTTTGTAGAAATGGGTTCGTTCTGTGAAGAG	1528
QY	1326	TATGACATTTGAACCCAGAGCTCCCGAGTTTACTTTTTCGCAAGAAATGATTTACCTGAA	1385
DB	1529	TACAACTAGAGCTTAGGGCCCCGGAATGGTACTTTGTCTCAGAAATTTGATTTACTTTGAA	1588
QY	1386	GACAAAGTCCAGCCTTCAATTTGTTAAAGACCGCCGGCCATGAAGAGAAATATCAAGAA	1445
DB	1589	GACAAAGTTCAAACCTCATTTGTGAAGAACCGCCGGCCATGAAGAGAAATATCAAGAA	1648
QY	1446	TTTAAATCAGATTAATGCCCTAGTTTCTTAAGCATTTGAAGTCCCGGAGGAGGATGG	1505
DB	1649	TTCAAAGTTCGATCAATGSGTCTTGTAGCAAGGCACAAAAGTTCCCGGAGGAGGATGG	1708
QY	1506	ATCATGCAAGATGGCACACCATGGCCAGGAAACATATCCAGGATCATCTCGGATGAT	1565
DB	1709	ATCATGCAAGATGGTACACCTTTGGCTGGGAAACATATCTAGGAGCATCTCGGATGAT	1768
QY	1566	CAGTTCCTTGTGTGTCAGTGTGGCTTGTATCTGAGGGTAAATGAGCTCCCGCTTTA	1625
DB	1769	CAGTTCCTTGTGTGTCAGTGTGGGCTTGTGCGTTGAAGCAATGAACTTCTCGTTG	1828
QY	1626	GTTTATGTGTCTGTTGAAAAGGCTCTGGTTCAGCACCAACAAAGGCTGGTGGCCATG	1685
DB	1829	GTTTATGTGTCTGTTGAAAAGGCTCTGGATTCCAACATCAAGAAAGGCTGGTGGCCATG	1888
QY	1686	AATGCCCTTGTGTGTCAGTGTCCCTTACTTAATGACAAATACATGTAATCTTCAAT	1745
DB	1889	AATGCCCTTGTGTGTCAGTGTCCCTTACTTAATGAGCAATACATGTAATCTTCAAT	1948
QY	1746	TGTGATCACTACATCAACAAAGGCTCTGCGAGAAAGCTATGCTCTCTTAATGAT	1805
DB	1949	TGTGACCACTACATCAATTAATAGCAAGGCTCTTCGAGAAAGCTATGCTCTCTTAATGAC	2008
QY	1806	CCAAACCTTAGTCCGCAAGTCTGTATGTGCAAGTTCCCAAAAGTTTGTATGGGATTCAT	1865
DB	2009	CCAAACCTTAGGAAAGATGTCTGTATGTCCAAATTCCTCAGAGGTTTGTATGGTATTCAT	2068
QY	1866	AGGAATGATCCATATGCAACAGGAAACATCTCTTTTGTATATTAATCTGAGGGCCCT	1925
DB	2069	AGGAATGACCAATATGCAACAGGAAACATCTGTGTTCGATATTAATCTGAGAGGCTCT	2128
QY	1926	GACGCAATTCAGGACCAAGTTTATGTGGAACTGTTGTGTTTTCACAGAACAGCTATC	1985
DB	2129	GACGCAATTCAGGACCAAGTTTATGTGGAACTGTTGTGTTTTCACAGAACAGCTATC	2188
QY	1986	TATGGTTATGAGCCCCCAATTAAGGCGGAAAGGAGCTCAGATAGAAAAGTTCGAAACAG	2045
DB	2189	TATGGTTATGAGCCCTCAGTCAAGAAAAGGAGGAGGCTCTCTCTCTCTCTCTCTCTCT	2248
QY	2046	GCGAGAGAGGCAACCAAGTCAAGAAAGGAGCTCAGATAGAAAAGTTCGAAACAG	2105
DB	2249	GGAAGGAAAAGAGCTCAAAATCTAAGAA--GAGCTCGGAAAAGAGAGTCAATAGA	2305
QY	2106	CATGTGACAGTCTCTGTTTCCAGTATTCATCTCCAGACATAGAGGAGGTTGTTGAAGT	2165
DB	2306	CACGACAGAGTCTCTGATCCAGTATTTAATCTCGAGATATAGAGGAGGATTTGAAGT	2365

[illegible]

QY	1423	CCATGAAGAGAGATATATCAAGAAATTTAAATACAGATAAATGCCCTAGTTCTTCTAAGGCAT	1482
Db	2130	CCATGAAGAGGATATAGAGAAATTTAAAGTTCGTGTTAATGCTCTCGTCGCCAAGGCAC	2071
QY	1483	TGAAAGTCCCGAGGAAGATGGATCATGCAAGATGGCACACCATGGCCAGAGAAACAATA	1542
Db	2070	AGAAAGTTCCCTGAGSAAGATGATATGCAAGATGGCACACCTTGGCTTGTAACAATA	2011
QY	1543	CCAGGATCATCTGGAAATGATTCAGGTTTTCTTGTCACAGTGGTGGCTTGATACTG	1602
Db	2010	CCAGAGACCATCTCGGAATGATTCAGGTTTTCTTGTCATATGTGAGAGGCTTGATACCG	1951
QY	1603	AGGTAATAGCTCCCGCTTTAGTTATGTGTCTCGTGAAAAGCGTCTGGGTTCCAGC	1662
Db	1950	AAGCAATAGCTTCCCTGTTTAGTCTATGTGTCTCGTGAAAAGCTCTCGGATTCACAC	1891
QY	1663	ACCAAGAAGGCTGGTCCCATGAATGCCCTTGTTTCGTGTCTCAGCTGTCTCTACTAATG	1722
Db	1890	ACCATAAAAGGCTGGTGCTATGAATGCACCTTGTTCTGTGTATCAGCTGTCTCTACTAATG	1831
QY	1723	GACAAATACATGTTGAATCTTGATTCATCATCATCAACAAGCAAGGCTGTCCGAG	1782
Db	1830	GACAAATCTGTTGAAATCTTGACGTGTATCATCATCATCAACAATAGCAAAAGCTCTCCGAG	1771
QY	1783	AAGCTATGTGCTTCCCTAATGGATCCAAACCTTAGTTCGCGAAGCTGTGTATGTGCGAGTTCC	1842
Db	1770	AGGCTATGTGCTTCCCTATGGATCCAAACCTTAGAAGGCTGTGTATGTGCAATTC	1711
QY	1843	CACAAAGTTTGATGGGATTGATAGGAATGATCGATATGCAAAACAGGAAACACTGTCTTTT	1902
Db	1710	CTCAGAGTTTGACGGTATCGATAGAAATGATCGATATGCAAAACAGGAAACACCGTGTTTT	1651
QY	1903	TTGATATTAATCTGAGGGCTTCACGGCATTCAAGGACAGTTTATGTGGAACTGTGTT	1962
Db	1650	TCGATATTAATTTGAGAGGCTCGATGCTCTCCAAGGACCAAGTTTATGTGGAAACCGGTT	1591
QY	1963	GTGTTTTCAACAGAACAGCTATCTATGTTTATGAGCCCCAAATTAAGGGCGAAGAGCCAG	2022
Db	1590	GTGTGTTCAACAGAACAGCTTTTATGTTTATGAAACCCCCCATTTAAGCAAGAGGCCAG	1531
QY	2023	GTPTCTTGGCATCATATGTGGGGCGAAGAGGCAAGCAAGTCAAGAAAAGGAGCT	2082
Db	1530	GTACTTCTCTTCGTTTGTGGGGGACGAAAGAAAGACAAAAGAGTCGAGGAGAGACA	1471
QY	2083	CAGATAGAAAAAGTTCGAAACAGCATGTGCACAGTTCTGTTCCAGTATTCATATCTCGAAG	2142
Db	1470	CGGAAAGAAAAAGTTCACAAAATGTGCACAGTTCTGTGCCAGTTTTTAATCTTTGAAG	1411
QY	2143	ACATAGAGGAGGTGTGTAAGAGTGTCTGGGTTTGATGATGAGAAATCAGTTCTCATGTCTC	2202
Db	1410	ATATAGAGGAAGGATTTGAGGTTCTGGAATTTGATGATGAGAAATCATCTAGATGCTC	1351
QY	2203	AAATGAGCTTAGAGAGAGATTTTGGCCAGTCAGCAGCATTTGTGCTCCACTCTCATGG	2262
Db	1350	AAATGAGCTTAGAGAAAAAGTTTGGTCAATCTAGTGTGTTTGTAGCCTCCACTCTCATGG	1291
QY	2263	AAATGTTGTTTCCCTCAGTCTCCACTCCAGAAATCTCTTTTGAAGAGAGCTATCCATG	2322
Db	1290	AAATGTTGTTTCCCTCAATCCGAACTCCGAAATCTCTTCTGAAAGAGGCCATACATG	1231
QY	2323	TCATAAGTTGTGCTATGAGGACAAGTCTCAATGGGGAACGTAGATTTGGTTGGATCTATG	2382
Db	1230	TTATCAGCTGTGCTATGAGACAAAAGTCACTGGGGAACTGAGATTTGATGATCTATG	1171
QY	2383	GATCTGTCAAGAGATATTTCTTACTGGATTCAGATTCAGCAAGAGGCTGCGGTTCCAG	2442
Db	1170	GTTCGTGTTACAGAAATATTTCTCACTGGATTCAGATTCAGATTCAGTGGCTGCGGTCAA	1111
QY	2443	TCATATGCATGCCAAGCGCCAGCTTTCAAGGGATCTGCCCCCACTCAATCTTTAGATC	2502
Db	1110	TCATCTGCATGCTTAAGCGACCAAGATTCAGGGGTTCTGCCCAATTAATCTTTCCGATC	1051

Search completed: August 23, 2004, 11:47:51
Job time : 1540 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2004, 01:15:51 ; Search time 8675 Seconds
(without alignments)
12481.875 Million cell updates/sec

Title: US-09-900-237A-29

Perfect score: 3626

Sequence: 1 gaacgaggaacccgcgtocac.....tcacattttggaggagtttt 3626

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST: *
1: em estba: *
2: em esthum: *
3: em estin: *
4: em estmu: *
5: em estov: *
6: em estpl: *
7: em estro: *
8: em htc: *
9: gb est1: *
10: gb est2: *
11: gb htc: *
12: gb est3: *
13: gb est4: *
14: gb est5: *
15: em estfun: *
16: em estom: *
17: em gss hum: *
18: em gss inv: *
19: em gss pln: *
20: em gss vrt: *
21: em gss fun: *
22: em gss nam: *
23: em gss mus: *
24: em gss pro: *
25: em gss rod: *
26: em gss phg: *
27: em gss vrl: *
28: gb gss1: *
29: gb gss2: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2572.2	70.9	3763	11	AY108113	AY108113 Zea mays
2	2402	66.2	3728	11	AY112236	AY112236 Zea mays
3	2199	60.6	3696	11	AY110079	AY110079 Zea mays
4	1373.2	37.9	3898	11	AY110415	AY110415 Zea mays

5	1320.4	36.4	2872	11	AY104236	Zea mays
6	1319.8	36.4	3788	11	AY103701	Zea mays
7	1315.6	36.3	3893	11	AY104730	Zea mays
8	1267.8	35.0	3787	11	AY103655	Zea mays
9	882	24.3	1874	14	CD726831	EST027 Cu
10	762.2	21.0	880	12	BM816138	BM816138 HCl09B12
11	751.6	20.7	916	12	BM816139	BM816139 HCl05G09
12	750.4	20.7	762	13	BQ802778	BQ802778 WHE2829 H
13	729.4	20.1	870	12	BG368813	BG368813 HVSMB1002
14	697.8	19.2	723	14	CD930401	GR45.111C
15	689.8	19.0	701	14	CD930401	GR45.111C
16	689.6	19.0	767	13	BQ578769	WHE0309 E
17	689.4	19.0	710	13	BQ282057	WHE3028 D
18	687.8	19.0	750	10	BF624748	HVSME001
19	685	18.9	706	14	CD877344	AZ04.100C
20	684.2	18.9	712	13	BQ620630	TaLR1138D
21	683.8	18.9	978	12	BG321258	Zm04.03f1
22	682	18.8	702	14	CD894946	G118.127I
23	670.4	18.5	865	14	CD440398	EL01N0554
24	669.4	18.5	718	14	CA501842	WHE4039 D
25	668.6	18.4	755	10	BF259426	HVSMEF001
26	668	18.4	694	14	CD880265	F1.001D08
27	664.2	18.3	680	14	CD878619	AZ04.103D
28	664.2	18.3	966	14	CA247044	SCC0AM209
29	663	18.3	705	9	AJ433167	AJ433167 AJ433167
30	651.4	18.0	659	12	BQ214460	BQ214460 BJ214460
31	650.4	17.9	702	14	CD907188	G468.106D
32	649.8	17.9	874	14	CD433620	EL01N0313
33	649.6	17.9	678	9	AJ433236	AJ433236 AJ433236
34	648.6	17.9	882	14	CK157933	FGAS03913
35	647.6	17.9	872	14	CK155617	FGAS03845
36	644.4	17.8	659	13	BQ620294	TaLR1173A
37	638.6	17.6	905	13	CA159025	SCE2R2312
38	637.6	17.6	676	13	BQ468579	HM01K12T
39	633.2	17.5	700	14	CB882206	HL01D01w
40	631.4	17.4	813	14	CK199867	FGAS00837
41	630	17.4	654	9	AV914933	AV914933 AV914933
42	628.4	17.3	661	12	BJ277746	BJ277746
43	628	17.3	656	13	BQ620013	TaLR1138D
44	625.4	17.2	671	14	CB868804	HC09H06w
45	625.2	17.2	646	14	CA500605	WHE4022_A

ALIGNMENTS

RESULT 1	AY108113	3763 bp	mRNA	linear	HTC 16-OCT-2002
LOCUS	AY108113	Zea mays	PCO126465	mRNA sequence.	
DEFINITION	AY108113	Zea mays	PCO126465	mRNA sequence.	
ACCESSION	AY108113				
VERSION	AY108113.1	GI:21211191			
KEYWORDS	HTC.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	1 (bases 1 to 3763)				
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.				
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes				
JOURNAL	Unpublished (2002)				
REFERENCE	2 (bases 1 to 3763)				
AUTHORS	Coe,E.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA				
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the				


```
QY 1864 ATAGGATGATCGATATGCAACAGCAACACTGCTCTTTTGTATATTAATTAATGAGGGCC 1923
Db 2136 ACAGGAATGATCGATATGCAACAGCAACACTGCTCTTTTGTATATTAATTAATGAGGGTC 2195
QY 1924 TTGACGGCATTTCAAGGACCAAGTTATGTTGGAACTGTTGTTTCAACAGAAACAGTAA 1983
Db 2196 TTGATGGCATCCAAGGACCAAGTTATGTTGGAACTGTTGTTTCAACAGAAACAGTAA 2255
QY 1984 TCTATGGTTATGAGCCGCCAAATTAAGCGAAAGCCAGGTTCTTGGCATCACTATGTG 2043
Db 2256 TATATGGTTATGAGCCGCCAAATTAAGCGAAAGCCAGGTTCTTGTGTCATCACTATGTG 2315
QY 2044 GGGCAAGAAAGGAGCAAGCAAGTCAAGAAAGAGGCTCAGATAGAAAGAGTCAACA 2103
Db 2316 GGGTGAAGAAAGGAGCAAGCAAGTCAAGAAAGAGGCTCAGATAGAAAGAGTCAACA 2372
QY 2104 AGCATGTGACAGTCTCTGTTCCAGTATTCATCTCGAAGACATAGAGGAGGTTGTTGAAG 2163
Db 2373 AGCATGTGACAGTCTCTGTTCCAGTATTCATCTCGAAGACATAGAGGAGGTTGTTGAAG 2432
QY 2164 GTGCTGGTTTCAATGATGAATCAGTTCTCATGTCTCAATGCTTAGAAGAGAT 2223
Db 2433 GGGCTGGATTTGACGACGAAATCACTTCTATGTCTCAATGCTTAGAAGAGAT 2492
QY 2224 TTGGCCAGTCAGCAGCATTTGTTCCCTCCACTCTGATGGAATATGTTGTTCTCACT 2283
Db 2493 TTGGCCAGTCAGCAGCATTTGTTCCCTCCACTCTGATGGAATATGTTGTTCTCACT 2552
QY 2284 CTTCACTCCAGATCTCTTTGAAGAGCTATCATGTCATAGTTGTTGCTATGAG 2343
Db 2553 CCGCAACTCCGAGTCTCTTCAAGAGCTATCATGTTATAGTCTGTCATGAG 2612
QY 2344 ACAAGTCTGAATGGGAACTGAGTTGGTTGGATCTATGATCTGTCAGAAATATTC 2403
Db 2613 ACAAGACTGAATGGGAACTGAGTCCGGTGGATCTACGGTCTGTGACAGAAACATTC 2672
QY 2404 TTACTGGATTCAAGTACGCAAGAGGCTGGGTTCACTCTATTCATGCTCCAGCGCC 2463
Db 2673 TCACCGGATTCAGATGACGCGGAGGCTGGGTTCACTCTATGCTGCTCCAGCGCC 2732
QY 2464 CAGCTTTCAAGGGATCTGCCCCCATCAATCTTTCAAGTCTGTAACCAAGTCTGCGGT 2523
Db 2733 CAGCTTTCAAGGGATCTGCCCCCATCAATCTTTCAAGTCTGTAACCAAGTCTGCGGT 2792
QY 2524 GGGCTCTGGTTCTGTTGAATCTTTTCAAGCGGATTTGCCCTTATGGTATGCTAGC 2583
Db 2793 GGGCTCTGGTTCTGTTGAATCTTTTCAAGCGGATTTGCCCTTATGGTATGCTAGC 2852
QY 2584 GAGGCGCTCAAGTTCTGGAGAGATTCGTTTACATCAACACCACTTTACCCACTAA 2643
Db 2853 GAGGCGCTCAAGTTCTGGAGAGATTCGTTTACATCAACACCACTTTACCCACTAA 2912
QY 2644 CTTCTCTCCGCTTCTAGTCTATTGTTATTTGCTGCTATCTGCTGCTCACTGGAAGT 2703
Db 2913 CTTCTCTCCGCTTCTAGTCTATTGTTATTTGCTGCTATCTGCTGCTCACTGGAAGT 2972
QY 2704 TCATCATGCCAGATTTAGCAACTTTGGCCAGTATCTGTTGTTATGTTGCTCTTTCAA 2763
Db 2973 TCATCATGCCAGATTTAGCAACTTTGGCCAGTATCTGTTGTTATGTTGCTCTTTCAA 3032
QY 2764 TTTTTCGCACTCGTATCTTGAAGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2823
Db 3033 TTTTTCGCACTCGTATCTTGAAGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3092
QY 2824 GGAATGAACAGTTCTGGGTCAATGGAGTATCTCTGCAATCTGTTTGGCGTCTTTCAAG 2883
Db 3093 GGAATGAACAGTTCTGGGTCAATGGAGTATCTCTGCAATCTGTTTGGCGTCTTTCAAG 3152
QY 2884 GTCTTCTGAGGTTCTGGGTATCGAACAACCTTCACTGCTCACTCAAGGCTAATG 2943
Db 3153 GTCTTCTGAGGTTCTGGGTATCGAACAACCTTCACTGCTCACTCAAGGCTAATG 3212
```

```
QY 2944 ACGAAGAGCGGAGCTTTGCTGAGCTCTACATGTTCAAGTGGAGCAGCGTCTTCTATCCCTC 3003
Db 3213 ACGAAGAGCGGAGCTTTGCTGAGCTCTACATGTTCAAGTGGAGCAGCGTCTTCTATCCCTC 3272
QY 3004 CAGAGGACCAATTTTATGATCAATTAACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3063
Db 3273 CACACCAATCTGATCATCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3332
QY 3064 ACAGTGGTTTACCAATCATGCGGCGCTCTTTTGGGAAAGCTCTTTTGGCTTCTGGGTGA 3123
Db 3333 ACAGCGGATACAGTCTGTTGGGCGCTCTCTCGCAAGCTCTTTTGGCTTCTGGGTGA 3392
QY 3124 TTGTTCACTTATACCAATCTTCAAGGTTTATGGGAGGCAAAACCGCACACCGACGA 3183
Db 3393 TCGTCCACCTGATACCGTTCTCAAGGGCTCATGGGAGGAGCAACCGCACACCGACGA 3452
QY 3184 TTGTCATCTCTGTTGGGCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3243
Db 3453 TCGTCTGTTGTTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3512
QY 3244 ATCCATCTACTACCGTCTGCTGCGCCAAATATCAAACTGTGGCATCAACTGCTAGG 3303
Db 3513 ACCCTTCAACCGCGCTCACTGCGCGGATACCCAGACGTGTGGCATCAACTGCTAGG 3572
QY 3304 ABAAGTGGAGTTT 3316
Db 3573 GAAGTGAAGGTT 3585
```

RESULT 2
AY112236
LOCUS CL1160_1 mRNA 3728 bp linear HTC 17-OCT-2002
DEFINITION Zea mays
ACCESSION AY112236
VERSION AY112236.1 GI:21216826
KEYWORDS HTC
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3728)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
Arthur,R.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 3728)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from zmdb and may be found by BLAST
searching at MSL, maizemap.org; Zmdb, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from Zmdb:
www.zmdb.iastate.edu.

FEATURES
Location/Qualifiers
1..3728
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630049"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"


```
QY 2136 CTCGAAGACATAGAGGAGGTGTTGAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTC 2195
Db 2311 NNNNNNNNTAGAGAGGGAGTTGAAGGCGTGGATTGATGATGAGAAATCAGTTCTT 2370
QY 2196 ATGTCTCAAAATGAGCTTAGAGAGAGATTGGCCAGTCAGCAGCAATTTGTTGCTCCACT 2255
Db 2371 ATGTCTCAAAATGAGCTTGAGAAGAGATTGGCCAAATCTGCAGCTTTGTTGGTCCACT 2430
QY 2256 CTGATGGAATATGGTGGTGTCTCTAGTCTCTCACTCCAGATCTCTTTTGAAGAAGCT 2315
Db 2431 CTGATGGAATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2490
QY 2316 ATCCATGTCTAAGTTGTGGCTATGAGGCAAGCTCTGAATGGGAACTGAGATTGCTG 2375
Db 2491 ATCCATGTCTAAGTTGTGGCTACGAGCAAGATTGAATGGGAACTGAGATTGCTG 2550
QY 2376 ATCTATGATCTGTCAAGAAGATATTCTTACTGGAATTCAGATGCACGCAAGAGCTG 2435
Db 2551 ATCTATGATCTGTGACGGAAGATATTCTACTTGGGTTCAAGATGCACGCAAGAGCTG 2610
QY 2436 CGTTCAGTCTATTGCAATGCCCAAGCGCCAGCTTTCAAGGATCTGCCCGCATCAATCTT 2495
Db 2611 CGTTCAGTCTATTGCAATGCCCAAGCGCCAGCTTTCAAGGATCTGCCCGCATCAATCTT 2670
QY 2496 TCAGATCGTCTGAACCAAGTGTGCGGTGGGCTCTCGTTCTGTTGAAATCTTTTCAGC 2555
Db 2671 TCAGACCGTCTGAACCAAGTGTGCGGTGGGCTCTCGTTCTGTTGAAATCTTTTCAGC 2730
QY 2556 CGGATTTGCCCTTATGATGATGCTACGAGGCGGCTCAAGTTCCTGGAGAGATTCGCT 2615
Db 2731 CGGATTTGNNNNATGATGATGATGCTACGAGGAGCGCTCAAGTTCCTGGAGAGATTCGCT 2790
QY 2616 TACATCAACACCAATTTACCCTTAACCTCTCTCCGCTCTAGTCTATTGTTATATG 2675
Db 2791 TACATCAACACCAATTTACCCTTAACCTCTCTCCGCTCTAGTCTATTGTTATATG 2850
QY 2676 CTTGCTATCTGCTCTCACTGGAAGTTTCATGCGCAGAGATTAGCAACTTGGCCAGT 2735
Db 2851 CTTGCAATCTGCTCTACGCGGAGTTTCATGCTCCAGAGATCAGCAACTTCGCTAGT 2910
QY 2736 ATCTGTTTCAATGCGCTCTCTCTTCAATTTTGGCACTGGTATCTTGGAGATGAGTGG 2795
Db 2911 ATCTGTTTCAATGCGCTCTCTCTTCTATCTCGATCTTTCGCAAGGTTATCTGAGATGAGTGG 2970
QY 2796 AGTGTGTGTCATTCAGAGTGTGGAGATGAACAGTTCTGGGTCTATGGAGGTATC 2855
Db 2971 AGCGGCGTGGGATCGACAGTGTGGAGGAAACGAGCAGTTCTGGGTCTATCGGAGGCATC 3030
QY 2856 TCTGCACATCTCTTTGCCGCTCTTTCAGGCTCTTCTGAAGGTCTTCCGCGTATCGACACC 2915
Db 3031 TCCGCCACCTCTTCCGCGCTTTCAGGCGCTCTTCAAGGTCTTCCGCGATCGACACC 3090
QY 2916 AACTTCACTGTCACTCAAGGCTAATGACGAAGGCGACTTTTCTGAGCTCTACATG 2975
Db 3091 AACTTCACTGTCACTCAAGGCTCGGATGAAGACGCGGACTTCCGCGAGCTGTACATG 3150
QY 2976 TTTCAAGTGAAGACGCTTCTATCCCTCGACGACCAATTTTCATCAATTAAGTGTGTG 3035
Db 3151 TTTCAAGTGAAGACGCTTCTATCCCGCCACCAACCATCTCTGATCATCAACCTGTGCGC 3210
QY 3036 GTCTGTGTGGCAGCTCTCTACGCCATCAACAGTGGTTTACCAATCATGGGGCGGCTCTTT 3095
Db 3211 GTTGTTCGGGATCTTCTATCGCCATCAACAGCGGGTACAGTGTGGGTTCGCTCTTTC 3270
QY 3096 GGGAGCTCTTCTTGGCCCTTCTGGGTGATTTGTTCACTTATACCCATTCCTCAAGGCTCTT 3155
Db 3271 GGGAGCTCTTCTTGGCCCTTCTGGGTGATTCGTTACCTGTACCCGTTCTCAAGGCTCTC 3330
QY 3156 ATGGGCGAGCAAAACGGACACGAGAGATTGTCATGCTCTGGGCTGTCTCTCTCGCTTCT 3215
Db 3331 ATGGGTCGGGCAAAACGGACACCCGACCATCGTGGTGTCTTGGGCGATCTGTCGGCGTCG 3390
QY 3216 ATCTTCTCTCTGTTGGTTCGTTGTTGATCCATTCACCTACCCGCTCTCGTGGCCCAAT 3275
```

```
Db 3391 ATCTTCTCTCTGTTGGGTTTCGATCGATTCGTTTCCCAACCGCTCTACTGCCCCGAT 3450
QY 3276 ATCCAAACCTGTGGCATCAACTGCTAGGAAAGTGGCA-GTTTGTAGAGACAGAAA 3329
Db 3451 ACTCGAAACGTGTGGCATCAACTGCTAGGAGGTGGAAAGTTTGTAGAARACAGAGA 3505

RESULT 3
AY110079
LOCUS Zea mays CL1164_1 mRNA sequence. 3696 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays
ACCESSION AY110079
VERSION AY110079.1 GI:21214162
SOURCE HTC.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizeimap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
Location/Qualifiers
1..3696
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630055"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Query Match 60.6%; Score 2199; DB 11; Length 3696;
Best Local Similarity 78.7%; Pred. No. 0;
Matches 2549; Conservative 0; Mismatches 681; Indels 9; Gaps 3;

QY 66 GCGACGCGGACGCGCTCAAGTCCGGGAGGACGCGGCGCGGAGCTGTGCCAGATCTGC 125
Db 257 GACGGCGGCGACGCCACCAATTCGGGAGAGCATGTGGCGGGGAGTGTGCCAGATCTGC 316
QY 126 GCGACGCGCTGGGACCACTGTTGACCGGCGAGCTTTCACCGCTCGAGCTGTGCCGC 185
Db 317 GCGACGCGGTGGGACCGCGCGGACGCGACCTCTTCCCGCTCGAGCTGTGCCGC 376
QY 186 TTCCCGTCTCCGCGCTGTACGACGACGAGGAGGAGGACCCAGGCTGTCCCTC 245
Db 377 TTCCCGTGTGCGCGCCATGCTACGAGTACGAGCAAGGACGCGACCGCTGTGCCGC 436
QY 246 CAGTCAAGACCAAGTACAAGCGCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 305
Db 437 CAGTCAAGACTAGTACAAGCGCCACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 496
```

Qy	306	GACGACACTGATCCGATGATGAGTGTGAGCTTCAACTACCTACCTGCATCTGSCACTGAGGAC	365	Qy	1386	GACAAAGTCAGGCTTTCATTTGTTAAAGACCGCGGGCCATGAAGAGAGAAATATGAAGAA	1445
Db	497	GAGGATGTGATCTGACGATGTGAGTGACTACAATACCAAGCATCTGGCAACACAGGAT	556	Db	1571	GACAAAGTTCAAACCTCATTTGTGAAGAAACGCGGNNNNNNNNNNNNNNNNNNNNNNNN	1630
Qy	366	CAGAAGCAGAAAGATTGCTGACAGATGCGCAGCTGGCGCATGAACACCGGGGCGAGTGGC	425	Qy	1446	TTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCATTGAAGTCCCGGAGGAGGATGG	1505
Db	557	CAGAAGCAAAAGATTGCTGAGAGAAATGCTCACTTGGCGGCAAACTCACGTGGCAGTG--	614	Db	1631	NNCAAAGTTTCGATCAATGCTCTGTAGCCAAAGGCCAENNNNNNNNNNNNNNNNNNNNN	1690
Qy	426	AATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTCTCCAAAGTATGACAGTGA	485	Qy	1506	ATCATCAGATGGCACACCATGGCCAGGAAACAAATACCAAGGATCATCTCTGGAATGATT	1565
Db	615	-AATATTGGCCTGGCTGAAGTATGACAGCGGTGAATTTGGGCATGGGAAGTATGACAGTGT	673	Db	1691	ATCATGCAAGATGGTACACCTTGGCCTGGGAACAATACATAGGACCAATCTCTGGAATGATT	1750
Qy	486	GAGATCCCTAGGAGTACGTCCTTCAGTCAACCAAGCCAGATGTCAGGAGAAATCCCT	545	Qy	1566	CAGGTTTCCCTCGTCCACAGTGTGGCTTGTATCTAGGGTAAATGAGTCCCCCGTTTA	1625
Db	674	GAGATCCCTCGTGATATATCCCGTCACTAACTCATAGCCAGATCTCAGGAGAGATTCCT	733	Db	1751	CAGGTTTCCCTGGGTACAGTGGAGGGCTTTCAGGTTGAAGCAATGAATCTCTCGTTG	1810
Qy	546	GGAGCTTCGCTCATCATCATGATGTCCTCCCTACGGGGAACATCAGCAACGTCGTCGG	605	Qy	1626	GTTTATGTGCTCGTGAAAGCGTCTCTGGGTTCCAGCACCAAGAAAGGCTGGTCCCATG	1685
Db	734	GGAGCTTCCCTCATCATATGATGTCCTCTGTTGGGAACATTTGGCAGGCGTGGACATCAA	793	Db	1811	GTTTATGTGCTCGTGNNNNNCGTCTGGATTCCAACATCACAAGAAAGGCTGGTCCCATG	1870
Qy	606	TTTCCCTATGTGAATCATTCACCAATTCCTCAAGGAGTTCTCCGCGAGTATTTGGGAAT	665	Qy	1686	AATGCCCTTGTCTCGTGTCTCAGCTGCTTACTTAATGGAACAATACATGTTGAATCTTGAT	1745
Db	794	TTTCCCTATGTAATCATTTCCAAACCCATCGAGGGAGTTCTCCGCTAGCCCTTGGCAAT	853	Db	1871	AATGCACTTGTTCGTGTATCAGCTGCTTACTTAATGGGCAATACATGTTGAATCTTGAT	1930
Qy	666	GTTGCTCGGAAAGAGAGATTGATGGCTGGAATGAAGCAGGACAAAGGCTGGGATTCCT	725	Qy	1746	TGTGATCACTPACATCAACAAAGCAAGGCTGTCGGAAGACTATGTCTTCCCTAATGGAT	1805
Db	854	GTTGCAATGGAAGAGAGGTTGATGGATGGAATGAATGA--AGATAAAGGTGCAATTCCT	910	Db	1931	TGTGACCACTPACATCAATATAGCAAGGCTCTTCGAGAAGCTATGTCTTCCCTATGGAC	1990
Qy	726	ATGACTAATGGGCAAGCAATGTCTCCCTGAAAGTCCGGCAGCTACTGATCGATGCA	785	Qy	1806	CCAAACCTAGTCCGCAAGTCTGTTATGTGCAGTCTCCCAAAAGGTTGATGGATGAT	1865
Db	911	ATGACCAATGGAACAAGCAATGTCTCCATCAGAAGGGCTGGAGTTGCTGATATTGATGCT	970	Db	1991	CCAAACCTAGGAAGGAATGTCTGTTATGTCCAAATTCCTCAGAGGTTGATGGATGAT	2050
Qy	786	TCTACTGAATACAACTGGAAGACGCTTACTGATGATGAATCGCAGCTCCGAGCCTCTATCT	845	Qy	1866	AGGAATGATCGATATGCAAAACAGAACACACTGTCTTTTTCGATATTAACCTTCAGGGGCTT	1925
Db	971	TCTACTGATTATACATGGAAGATGGCTTACTGAATGATGAATCGGCAACCTCTATCT	1030	Db	2051	AGGAATGACCGATATGCAAAACAGAACACTGTGNNNNNCGATATTAACCTTCAGAGGCTT	2110
Qy	846	AGAAAGTCCCATTTGCTCTCCAAATAAATCCCTACAGAAATGGTCAATGTTCTCGCG	905	Qy	1926	GACGGCAATCAAGAACAGTTTATGTGGAACTGTTGTGTTTTCACAGAACAGCTATC	1985
Db	1031	AGAAAGTGCAATTCCTTATCCAGAAATAATCCGTACAGATGGTCAATGTTGCTACGT	1090	Db	2111	GACGGCAATCAAGGGCCAGTTTATGTGGAACTGTTGTGTTTTCACAGAACGGCCCTTA	2170
Qy	906	TTGGTTGTTCTTAAGCATCTTCTGCACTACCGTCTCAAAATCCTGTGCGTATGCAATAC	965	Qy	1986	TATGGTTATGAGCCCCCAATTAAGGCGAAGAACGAGTTCCTTGGCATCACTATGTGGG	2045
Db	1091	TTGGCTGTTCTATGCAATTTCTTGGCTACCGTATACACATCCTGTGAAACAATGCATAT	1150	Db	2171	TATGGTTATGAGCTCCAGTCAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	2230
Qy	966	CACTGTGGCTTTTATCTGTTATATGTGAGATTTGGTTTCTTATCCTGGATCTGGAT	1025	Qy	2046	GGCAAGAAAGCAAGCAAGTCAAGAAAGAGAGCTCAGATAGAAAAAGTTCGAAACAAG	2105
Db	1151	CCACTGTGGCTTTTATCCGTCATATGTGAGATCTGGTTTGTCTTGTCTGGATTTTGGAT	1210	Db	2231	NNAAGNNNNNAGCTCAAAATCTAAGAA--GAGCTCGGAAGAAGAGTCAATAGA	2287
Qy	1026	CAGTTCCGGAAGTGGTTTCCAAATCAACCGGAGACCTTACCTTGATAGACTGGCTTTAAGG	1085	Qy	2106	CATGTGGACAGTTCTGTTCAGTATTCAAATCTCGAAGACATAGAGGAGGTTGTAAGGT	2165
Db	1211	CAGTTCCCAAGTGTCCCNATCAACCGTGAACATACCTTGATAGACTGGCTTTAAGG	1270	Db	2288	CAGCAGACAGTTCTGTACCAGTATTTAAATCTCGAAGATATAGAGGAAGGATTTGAAGGT	2347
Qy	1086	TATGACCGAAGGTGAACCGTCTCAGTTGGCTGCTGTGACATATTTGTGACGATACGATC	1145	Qy	2166	GCTGGGTTTATGATGAGAAATCAGTTCTCATGCTCAAAATGAGTTCCTCAGTCC	2225
Db	1271	TATGACCGAAGGTGAACCATCTCAATTAGCTCCTGTTGATANNNNNGTCAGTACTGTG	1330	Db	2348	TCTCAGTTTATGATGAGAAATCGCTGATTTATGTCTCAAAATGAGCTTGGAGAAGATTT	2407
Qy	1146	GACCCCTTGAGAGGACCACTATCGTCACTGCCAACACTGTGCTATCCATCTTGTCTGTT	1205	Qy	2226	GGCCAGTCAGCAGCATTTGTTGCCCTCCACTCTGATGGAATATGGTGGTTCCTCAGTCC	2285
Db	1331	GATCCAATGAGGAGGCTCTCTTGTCACTGCAAAATACGTGCTTTCCATCCTTGTCTGTC	1390	Db	2408	GGCCAGTCCAGTGNNNNNNGTAGCTCTACTGATGGAATATGGTGGTGTCCACAATCT	2467
Qy	1206	GATTATCCCGTGGCAAGGCTCTCTGCTATGATGATGACGAGCTTCAATCCTGCTGACT	1265	Qy	2286	TCCACTCCAGAAATCTCTTTTGAAGAAGCTATTCATGTCTAATGCTATGGCTATGAGGAC	2345
Db	1391	GATTATCCGTTGACAGGATCTCTGTAATGTTTCGGATGATGAGCTGCTATGCTGACT	1450	Db	2468	GCAACTCCNN	2527
Qy	1266	TTTGACCGATTTGGCTGAGATTCAGATTTGCTAGGAAATGGGTACCATTTGTGGAAGAAG	1325	Qy	2346	AAAGTCTCAATGGGAACTGAGATTTGGATCTATGATCTGATCAAGAGATATTTCTT	2405
Db	1451	TTTGATGCTCTCTCTGAAATTCAGAGTTTGTAGAAAATGGGTTCCGGTCTCTGTAAGAAG	1510	Db	2528	AAAATGACTGGGGAATGAGATTTGGGTGGATCTATGTTCTGTTACAGAGACATCTC	2587
Qy	1326	TATGACATTTGAACCGAGGCTCCGAGTTTACTTTTCCAGAGAAATGATTACTGAAA	1385	Qy	2406	ACTGGATTAAGATGACGCAAGAGGCTGGCGTTTCACTTATTCATGCTCCCAAGCGGCCA	2465
Db	1511	TACAACATAGAGCCTANGGCCCCCGAATGGTACTTTGCTCGAAGAAATGATTACTTGAAA	1570	Db	2588	ACCGNN	2647
				Qy	2466	GCCTTCAAGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTCTCGGTGG	2525

QY	650	CGCAGCTATTGGGAATGTTGCTCGAAAGAGAGAGTTGATGCTCGAAAAATGAACAGGA	709
Db	782	CTATGGGCTTAATAGTGTGGACTGGAAGGAAGAGTTGAGAGCTGGAGGTTTAAACAGGA	841
QY	710	CAAGGGTGCATTTCCCATGACTTAATGGGCAAGCAATTGCTCCCTCTGAAGGTCGGGCAGC	769
Db	842	CNNNNNTATTGTCGAAGTGACTAATAAATATCCAGAGGCTAGAGGAGGAGACATGGAGGG	901
QY	770	TACTGCATCGATGCATCTACTGAATACAACATGGGAAGACGCTTTTACTCAATGATGAAC	829
Db	902	GACTGGCTCAAAATGGA-----GAAGATATGCAATGGTTGATGATGC	943
QY	830	TCGCCAGCCTCTATCTAGAAAAGTCCCATFTGCTCTCCATCCAAAATAAAATCCCTACAGAA	889
Db	944	ACGCTACCTTTGAGCCGATCGTGCCAAATTTCTCAAAACAGCTCAACCTTTACCGGT	1003
QY	890	GGTCATTGTTCTCGGTTGGTTGTTCTAAGCATCTTCTTGCACTACCGTCTCACAAATCC	949
Db	1004	AGTGATCATTTCTCGTCTTATCATCTCTGTGCTTCTTTCCAGTATCGTATCAGTCATCC	1063
QY	950	TGTGGCTAATGCATACCCATCTGTGGCTTTTATCTGTTATATATGAGATTTGGTTTGCTTT	1009
Db	1064	AGTCGGTGATGCTTATCGAATGTGGCTAGTATCTGTTATCTGCGAGGTCTGGTTTGCTTT	1123
QY	1010	ATCCTGGATACTCGATCAGTTCCCGAAGTGGTTTCCAAATCAACCGGAGACCTTACTTGA	1069
Db	1124	GTCTGGCTTCTAGATCAGTTTCCCAAATGGTATCCAATCAACCGTGAGACATATCTCGA	1183
QY	1070	TAGACTGGCTTTAAGGTATGACCGAAGGTGAACCGTCTCAGTTGGCTGCTGTTCACAT	1129
Db	1184	CAGGCTTGCATTTGAGGTATGATAGAGGGAGAGCCATCACGCTGGCTCCCATTTGATGT	1243
QY	1130	ATTTGTGACGTACGTCCGACCCCTTGAAGGAGCCACCTATCGTCACTGCCAACTGTGCT	1189
Db	1244	CTTGTGAGTACAGTGGATCCATTTGAAGGAACCTTCACTGATCAGGCCACACTGTTTT	1303
QY	1190	ATCCATTCTTGCTGTGATTATCCCTGGGCAAGGCTCTTGTGATGTATCTGATCAACGG	1249
Db	1304	GTCCATTCTTGCTGTGGATTACCTGTTGACAAAGTGTCTATGTTTCTGATGATGG	1363
QY	1250	AGCTTCAATGCTGACTTTTGACGCAATTTGGCTGAGACTTCAAGTTTGTCTAGGAAATGGCT	1309
Db	1364	CTCAGCTATGCTGACTTTTGAGTCTCTCTGTAACCGCGCAATTTGCTAGAAAGTGGGT	1423
QY	1310	ACCATTTGTGAGAGATGATGATTAACCCAGAGCTCCGAGTTTTTACTTTTTGCCAGAA	1369
Db	1424	TCCCTTTGTGAGAGACCAATATTTGAACCAAGAGCTCCAGAAATTTTACTTTGCTCNNN	1483
QY	1370	AATTGATTACCTGAAAGACAAAGTCCAGCCTTCAATTTGTTTAAAGACCGCGGCCATGAA	1429
Db	1484	NNTAGATTACCTGGAAGGACAAAATTCACCTTCAATTTGTTTAAAGAAAGACGCGCAATGA	1543
QY	1430	GAGAGAAATAGAAATTTTAAATTCAGGATAAATGCCCTAGTTTCTTAAGGCATTTGAAGT	1489
Db	1544	GAGGGAGTATGAAGAAATTCAAAGTAAGAAATCAATGGCCCTTGTGGCCAAAGCACAGAAAGT	1603
QY	1490	CCCCGAGGAAGGATCGGATCATGCAAGATGGCACACCATGGCCAGGAAACAATACCAGGA	1549
Db	1604	GCTTGAAGAGGGGTGGACCATGGCTGATGGAACCTGTGCTTGGCTGGGAAATACCCTAGGA	1663
QY	1550	TCATCTCGGAATGATTCAGGTTTTTCTTTGGTTCACAGTGGTGGCTTCATACTCAGGGTAA	1609
Db	1664	CAATCCTGGCATGATTCAGGTGTTCTTGGGACAGTGGTGGGCTCGACACTGATGAAA	1723
QY	1610	TGAGTCCCGCTTTAGTTTATGTGCTTCGTGAAAGCGTCTCTGGGTTCCAGCACACCAA	1669
Db	1724	TGAGTTAGCACGCTCTGTCTATGTCTCTCGTGAAGAGACCGAGCTTTTCCAGCATCAA	1783
QY	1670	GAAAGCTGGTCCATGAATGCCCTTTGTTTGGTTCACAGTGGTGGCTTCATACTCAGGGTAA	1729
Db	1784	GAAAGCTGGTGAATGAATGAATCGCTGATTCGTGTTATCTGCTGTGCTGACAAATGTTGCTTA	1843
QY	1730	CATGTTGAATCTTGATGTGATCACTACATCAACACAGCAAGGCTGTCCGGAAGCTAT	1789

1844	TCTTCTCAATGTGGATTTGGACCAATTACTTCAATAGCAGCAAGCTCTTAGAGACCAAT	1903
1790	GTGCTTCTCTAATGGATCCAAACCTAGGTCCGCAAGTCTGTTATGTGCAGTTCCTCCACAAAG	1849
1904	GTGCTTCATGATGGATCCGGCTCTAGGAAAGGAAACCTGTATGTACAATTTCCACAGAG	1963
1850	GTTTGTGGGAATTGATAGGAATGATCGATATCGATCGAATCGAAGAACACTGTCTTTTGTGATAT	1909
1964	AATTTGATGGCAATTGACTTGCACGATCGATATGCTAATCGGAACATAGTCTCTTTTGTGATAT	2023
1910	TAACTTGAGGGGCTTTGACGGCAATTCAGGACACAGTTTATGTGGAACTGGTGTGTGTTT	1969
2024	CAACATGAAAGGTCTGATGCGCATTCAGGGTCCAGTCTACGTGGGAAACAGGATGCTGTGTT	2083
1970	CAACAGAAACAGCTATCTATGTGTTATGAGCCCCCAATTAAGCGGAAGAACGCCAGGTTTCTT	2029
2084	CAATAGGCAAGCTTTGTATGGATACGA-TCTCTGTTTGTAGTGAAGCTGATCTGTGGAGCCATA	2142
2030	GGCATCACTATGTGGGGGCAAGAAGGCAAGCAAGTCAAGAAAAAGGAGCTCAGATAA	2089
2143	ACATTTGTTTGAAGCTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATATGATA	2202
2090	GAAAAAGTCAAAACAAAGCATGTGACAG- ---TTCGTTCAGTATTTCAATCTCGAAGACA	2145
2203	GTCAAGCCGTATTATTGAAGAGAACAGAACTTTCAGCTCCCACTTTCAACATGGAAGACA	2262
2146	TAGAGGAGGTGTGTGAAGTGCTGGGTTTGATCATGAGAAATCAGTTCTCATGCTCTCAAA	2205
2263	TCGAGGAGGGTATTGAAGGT-----TACGAGGATGAAGTCTGAGTCTTATGTCGCCAGA	2316
2206	TGAGCTTTAGAGAAGAGATTTGGCCAGTCTCAGCAGCATTTGTTGCCTCCACTCTGATGGAAT	2265
2317	GGAAATTTGGAGAAACGCTTTGGTCAGTCTCCAATCTCATTTGCATCCACCTTTATGACTC	2376
2266	ATGCTGTTGTCCTCAGTCTCTCCACTCCAGAACTCTTTTGAAGAGCTATCCATGTCA	2325
2377	AAGGTGCATACCACTTTCAACAAACCCGCTTCTCTACTAAAGGAGCTATCCATGTCA	2436
2326	TAACTTTGGCTATGAGCAACAAGTCTCAATGGGGAACTGAGATTTGGTTGATCTATGGAT	2385
2437	TCAGTTGTGGATAGAGACAAACCTGAATGGGGAAAGAGATTTGGCTGATCTATGGTT	2496
2386	CTGTCAAGAGATATTCTTACTGGAATCAAGATTCAGCAAGCAAGGCTGCGGTTCACTCT	2445
2497	CAGTTACGGAGGATTTCTGACTGGGTTTAAANNNNNNNNAGGGGCTGCCAATCAAATCT	2556
2446	ATTGCATGCCCAAGCGGCCAGCTTTCAAGGATCTGCCCCCATCAATCTTTTCAGATCGTC	2505
2557	ACTGCATGCCACACAGACTTGTTTCAAGGGTCTGCAACCAATCAATCTTCTGATCGTC	2616
2506	TGAACCAAGTCTCGGTGGGCTCTCGGTTCTGTTGAAATCTTTTTCAGCGGGCATTTGCC	2565
2617	TTAATCAGGTGCTCCGTTGGGCTCTTGGGTCAAGTGGAAATTTCTGCTTAGCAGACATTTGTC	2676
2566	CCTTATGATATGGCTACGGAGGGGCTCAGTTTCTGGAGAGATTCGCTTACATCAACA	2625
2677	CTATCTGATATGGTTTACAATGGAAGATTTGAAGCTTTTGGAGAGCTGGCTTACATCAACA	2736
2626	CCACCATTTACCACTAACCCTCTCTCCGGCTCTTAGTCTATTGTTATATTCCTGCTATCT	2685
2737	CCATTTGTTATCCAAATCACATCTGTTCCGCTTATCGCTATTGTTGTGCTTCTGCTATCT	2796
2686	GTCTGCTCACTGGAAAGTTCAATGACGAGATTAGCACTTGGCCAGTATCTGGTTCA	2745
2797	GTCTTCTTACCAATAAAATTTATCAATCTCGAGATTAGCAATATGCTGGGAGTGTCTTCA	2856
2746	TTGCGCTCTCTCTTTCAATTTTTGCCCACTGGTATCCCTTGAGATCAGGTGGAGTGGTGTG	2805
2857	TTCTTCTTTTTCGCTTCCANNNNNNCCATGCTATATTGGAGCTTTAGATGGAGTGGTGTG	2916
2806	GCATTTGACAGGTGGTGGAGGAATGAAAGATTCTGGGTCAATTTGGAGGTATCTCTGCACATC	2865

```

Db      2917 GCATTGAAGATTGGTGAGAAATGAGCAGTTTGGGTTATTGGTGCCACCTCTGCCCATC 2976
QY      2866 TGTTCCTCCCTCTTTCAGGGTCTTCTGAAGTGCTTGCCGTATCGACACCACTTCACTG 2925
Db      2977 TCTTCGAGTGTCCAGGGTCTCTGAAAGTGTGGCTGGGATGATACCACTTCACAG 3036
QY      2926 TCACCTCAAAGGCTAATGACGAAGAAGGCGACTTTGCTGAGCTCTACATGTTCAAGTGGA 2985
Db      3037 TTACCTCAAAGGCATCTGATGAGGATGGGACITTTGCTGAGCTATATGTTCAAAGTGA 3096
QY      2986 CGAGCGTCTTCATCCCTCCGACGACCATTTTGAATCAATTAACATGGTGGTGGTGGCTG 3045
Db      3097 CAGTTTGCTCAATCTCTCCGACACATGTTCTTGTCAATTAACCTGGTCGGAATGGTGGCAG 3156
QY      3046 GCACCTCTACGCCATCAACAGTGGTTACAATCATGGGGCGCTTTGGGAAGCTCT 3105
Db      3157 GAATTTGGTATGCCATTAAACAGTGGCTACCAATCTGGGGTCCGCTCTTTGGAAAGCTGT 3216
QY      3106 TCTTTGGCTTCTGGGTGATGTTCACTATATACCATTCCTCAAGGCTCTTATGGGCGAGC 3165
Db      3217 TCTTCTCGATCTGGGTGNTCCCTCAATCTGTACAGCTTCTCAAGGCTCTCATGGGAAGC 3276
QY      3166 AAAACCGCACCGAGGATGTCATGCTCTGGGCTGCTCTCGCTTCTATCTTCTCT 3225
Db      3277 AGAACCGCACACCAACAAATCGTCAATGCTGTGCTGCTCCATCCTTCTTGATCTATCTTCTCT 3336
QY      3226 TGCTGTGGGTTCTGTTGATCCATTCACTACCC 3258
Db      3337 TGCTGTGGGTGAAGATCGATCTTTCATCTCCC 3369

RESULT 5
AY104236
LOCUS      Zea mays PC0121439 mRNA 2872 bp linear HTC 16-OCT-2002
DEFINITION
ACCESSION AY104236
VERSION   AY104236.1 GI:21207314
KEYWORDS  HTC.
SOURCE    Zea mays
ORGANISM  Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2872)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 2872)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizeimap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
    source
        1. .2872
            /organism="Zea mays"
            /mol_type="mRNA"
            /db_xref="MaizeDB:637064"
            /db_xref="taxon:4577"
            /clone_lib="Maize Mapping Project/DuPont Consensus
            Library"
            /note="this sequence is part of a project of EST
            assemblies resulting from the application of public
            contigs to seed DuPont contigs; this resource was
            assembled by DuPont as part of a collaboration for the

```

```

overgo addressing of BACs in conjunction with the Maize
Mapping Project"
Query Match      36.4%; Score 1320.4; DB 11; Length 2872;
Best Local Similarity 72.2%; Pred. No. 6.2e-265;
Matches 1760; Conservative 0; Mismatches 666; Indels 10; Gaps 3;

ORIGIN
QY      829 CTGCGCAGGCTCTATCTAGAAAAGTCCCATTTGCTTCTCTCCAAATAAATCCCTACAGAA 888
Db      14  CACGGCTACTCTAAGTCGATAGTTCCGATATCTCCAAACGAGCTTAACCTTTATCGGA 73
QY      889 TGGTCAATCTCTGCGGTTGGTTCTTAAGCATCTTCTGCACTACCGTCTCAAAATC 948
Db      74  TCGTGAATGTTCTCGGGCTTATCATCTATGTTTCTTCTTCAATATCGTAACTCATC 133
QY      949 CTGTCGTAAATGCATACCCACTGTGGCTTTTATCTGTATATCTGAGATTGTTGGTTGCTT 1008
Db      134 CAGTGGAGATGCTTATGGGTTGTGGCTTGAATCTGTTATTTGTGAAGTTTGGTTGGCT 193
QY      1009 TATCCTGGATATCTGGATCAGTTCCCGAAGTGGTTTCCAAATCAACCGGGAGACCTACTTGG 1068
Db      194  TGTCTTGGCTTCTAGATCAGTTCCCAAGTGGTATCTTATCAACCGTGAACCTTACTCTG 253
QY      1069 ATAGACTGGCTTAAAGTATGACCGAGAAGGTGAACCGTCTCAGTT-GGCTGCTGTGAC 1127
Db      254  ATAGACTTGCATTTGAGATGATGATGGGAGGGTGAAGCCATCCAGTTGGGCTCCAATCGAT 313
QY      1128 ATATTGTGCAGTACAGTGCACCCCTTGAAGGAGCCACCTATCGTCACTGCCAACACTGTG 1187
Db      314  GTCTTTGTAGTACAGTGGATCCACTTAAGNAACCTCTCTAATTACTGGCAACTGTG 373
QY      1188 CTATCCATCTTGTCTGTTGATTATCCCTGGACAAGTCTCTTGTGTATGTATCTCATGAC 1247
Db      374  CTGTCCATCTTGTCTGTGGATTACCTCTGTGACAAGTATCATGTTAATGTTTCTCATGAC 433
QY      1248 GGAGCTTCAATGCTGACCTTTTGAACCATTTGGTGCATTTGAGAGTTTGTAGGAAATGG 1307
Db      434  GGTTCAGCTATGTTGACTTTTGAAGCGCTATCTGAAACCGCAGAGTTTTCGAAGGAATGG 493
QY      1308 GTACCATTTGTGAAGAAGTATGACATTGAACCCAGAGCTCCCGAGTCTTCTTTTGGCCAG 1367
Db      494  GTTCCCTTTTCGACGAACAACAATATGAACCTAGGGCTCCAGAGTTTACTTTTCTCTCGA 553
QY      1368 AAAATTGATTACCTGAAAGACAAAGTCCAGCTTCAATTTGTTAAAGACCGCGGGCCCATG 1427
Db      554  AAGATAGATTACCTAAAGGACAAAATAACAACCTCTTTTGTGAAGAAGGCGGCTATG 613
QY      1428 AAGAGAGATATGAAGAATTTAAATCAGGATTAATGCCCTTAGTCTTCTAAGGCAATGAAA 1487
Db      614  AAGAGGGAGTGTGAAGAGTTCAAAGTACGGATCGATGCCCTTGTTCAAAAGCGCAAAA 673
QY      1488 GTCCCCGAGGAAGGATGGATCATCGAAGATGSCACCATCGCCAGGAAACAATACCAAG 1547
Db      674  ATACTGAGGAGGCTGGAACATGGCTGATGGCACTCTTGTGGCTTGGGAAATACCCCTAGA 733
QY      1548 GATCATCTCGAATGATTCAAGTTTTCCTTGGTCAACAGTGGTGGCTTGTACTACTGAGGT 1607
Db      734  GATCATCGGGAATGATCCAAAGTATTCTTGGGCCACAGTGGTGGCTTGACACCGATGGG 793
QY      1608 AATGAGCTCCCCGTTTAGTTTATGTTCTGCTGTAAGAGCGTCTCGGGTTCACGACCAAC 1667
Db      794  AATGAGTTGCCAGGGCTTGTATTGTTTCTGTGTAAGAGGCGGCGAGGCTTCCAGCACCA 853
QY      1668 AAGAAGGCTGGTGCATGAATGCCCTTGTTCGTGTCTCAGCTGTCTTACTATATGACAA 1727
Db      854  AAGAGGCTGGTGCATGAATGCTTTGATTCCGGTATCAGCTGTCTTCGACGAATGTGCT 913
QY      1728 TACATGTTGAATCTTGAATTTGATCACTACATCAACAAACAGCAGGCTGTCCGAGAGCT 1787
Db      914  TATCTTCTTAATGTGGATTTGATCACTACTTCAATAGCAGCAAGCTCTTAGAGAGCT 973
QY      1788 ATGTGCTTCTTAATGGATCCAAACCTAGTCCGCAAGTCTGTTATTGTGAGTCTCCCAAA 1847

```


2339 AGGTGCTCCAGTGC-----TGATATCGAAGAGCCGGAATCGTAAATCAACAGAAACT 2392
Db
2213 AGAAGAGATTGGCGCATCAGACGACATTTGTTGCTCCTCACTCTGATCGAATATGGTGG 2272
Qy
2393 AGAAGAGAAATTTGGCGCATCTCTCTTTTGTGCGCATCAACACATTCCTTGAGAACGGAG 2452
Db
2273 TGTTCCTCAGTCTCCACATCCAGAAATCTCTTTTGAAGAAGCTATCCATGTCTAATAGTTG 2332
Qy
2453 GACCTCGAAGAGCGCAAGTCCAGCTCTCTCTCTGAAGGAGACTATACATGTTATCAGCTG 2512
Db
2333 TGGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTGGTTGGATCTATGGATCTGTAC 2392
Qy
2513 CGGCTACGAGACAAAGACCGACTGGGMAAGAGATTGGCTGATTTACGATCGATCAC 2572
Db
2393 AGAAGATATCTTACTGGAATTCAGATTCAGATTCGCAAGAGGCTGCGGTTCACTTATTCGAT 2452
Qy
2573 AGAGGATATCTTACTGGAATTTAAGATTCGACTGCCATGGCTGGCGGTCTATTACTGCA 2632
Db
2453 GCCAAGCGCCAGCTTCAAGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCA 2512
Qy
2633 CCGAAGCGCCCTGCAATTCAGAGTTCTGCGCTCTGAACCTTCCGACGCTCTCACCA 2692
Db
2513 AGTGTCTCGGTGGGCTCTCGGTTCTGTGTAATTCCT--TTTCAGCGGCAATGCCCTTAT 2571
Qy
2693 GGTCTCTGCTGGGCCCTTGGGTCCGTGGAATTTCTTTTCAGNAANCAATGNCACATTT 2752
Db
2572 GGTATGCTACGAGGCGCTCAAGTTCCTGGAGATTCGCTTACATCAACACACCA 2631
Qy
2753 GGTACGATACGGCGGCGCTAAATTCCTGGAAGGTTTCTTATCAATCAATCCATCG 2812
Db
2632 TTTTACCACATCACTCTCTCCGCTTCTAGTCTATTTATATGATTCCTGCTATCTGCTGC 2691
Qy
2813 TTTTATCCTGGAGCTCAATTCCTCTCTGGCTTACTGACTTTCCTGCTGCCATCTGCTGC 2872
Db
2692 TCACCTGAAAGTTTCAATCATCCAGAGATTCAGCAATTCGCGAGATCTGTTGTTCAATCCG 2751
Qy
2873 TCACGGGAAGTTTATCACACAGAGCTTACCAATGTCGCGAGTATCTGTTTCATGTCAC 2932
Db
2752 TCTTCTTTCAATTTTCCGCACTGGTATCCTTGAGATGAGTGGAGTGTGTTGGCATG 2811
Qy
2933 TTTTTCATCTGCATCTCCGTGACCGGCATCTCTGGAATGAGTGGAGTGGCGTGGCCATCG 2992
Db
2812 ACAGTGTGAGGAATGAACAGTTCTGGGTTCATTTGGAGGTATCTCTGCACATCTGTTG 2871
Qy
2993 ACAGTGTGAGGAACAGCAGATTCTGGGTATCTCGAGGGGTTTCGGGCACTGTTGTCG 3052
Db
2872 CCGTCTTTTCAAGGTCTTTCTGAAGTGTCTGCGGTATCGACCAATTCATCTGTACCT 2931
Qy
3053 CGGTGTTCCAGGGCTGCTGAAGGTGTTGGCGGCATCGACACAGAGCTTCAACCGTGAGT 3112
Db
2932 CAAGGCTAATGACGAAGAGCGACTTGTGAGCTCTACATCTTCAAGTGGACGAGC 2991
Qy
3113 CGAAGCGGGGACGACGA---GGAGTTCTCGGAGCTGTACAGTTCAAGTGGACCAACC 3169
Db
2992 TTCTCATCCCTCCGACGACCAATTTTGATCATTAACATGTTGGTGTGCTGCTGCGACCT 3051
Qy
3170 TGTGATATCCCGACACAGCTCTCTGCTGCTGATCTCATCGGGTGTGTCGCGGATCT 3229
Db
3052 CCTACGCGCATCAACAGTGGTTTACCAATCATGGGGCCGCTCTTTGGGAAGCTCTTCTTTG 3111
Qy
3230 CGAACCGCATCAACAAACGGGTACGAGTCTGGGGCCCTCTTTCGGGAAGCTCTTCTTTCG 3289
Db
3112 CTTTCTGGGTATGTTTCACTTATACCATTCCTCAAGGCTTATTTGGGAGGCAAAACC 3171
Qy
3290 GTTCTGGGTGATCGTCACTGATCCCGTCTCTCAAGGGTCTGTTGGGGAGGCAAGACA 3349
Db
3172 GCACACCGCATGTCATCTGGGTGTCCTCTCTGCTTCTATCTTCTCTTCTGCTGT 3231
Qy
3350 GGAACCGCATCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3409
Db
3232 GGGTTCGTTGATCCATTCATACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3291
Qy
3410 GGGTCCGCTGACACCGTCTCTTCGCAAGAGCGACGGCCCGCTCTCTGAGGAGTGTGCCC 3469
Db

Qy 3292 TCAACTGCTAGGAAGTGGAG 3313
Db 3470 TGGACTGCAACTGAAGTGGGG 3491
RESULT 7
AV104730 Zea mays P0100501 mRNA 3783 bp linear HTC 16-OCT-2002
LOCUS Zea mays
DEFINITION Zea mays P0100501 mRNA sequence.
ACCESSION AY104730
VERSION AY104730.1 GI:21207808
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3783)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 3783)
Coe, E.H.
DIRECT SUBMISSION
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
LOCATION/QUALIFIERS
1. 3783
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635765"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Query Match 36.3%; Score 1315.6; DB 11; Length 3783;
Best Local Similarity 64.4%; Pred. No. 6.7e-264;
Matches 2132; Conservative 0; Mismatches 1074; Indels 102; Gaps 7;
Qy 47 GCTCTGCGCGCCATGACGCGGACGCGGACCCCTGAAGTCCGGAGGACCGGGCCGG 106
Db 233 GCTCGTCTCATCCGCGCATGCGAGCGGCGGCGAAGCCCATGGACCGGAAACGG 292
Qy 107 GGACGTGTGCCAGATCTGCGCGGACGCGCTGGGACACACGTTGGACGCGGACGTTCCAC 166
Db 293 CCAGGTGTGCCAGATTTGCGGCGACGAGCTGGGGCGCAACCCCGGAGGAGCCGTTCT 352
Qy 167 GCGCTGGGACGCTCTGCGCTTCCCGGTCTGCGGCGCTTACGAGCAGCGGCAAGGA 226
Db 353 GGCCTTGCAACAGATGGGCTTCCCACTCTGCGGAGACTGCTACGAGTACGAGCGCGGA 412
Qy 227 GGGCAACCCAGGCTCTGCTTCCAGTGCAGACCAAGTACGAGCGCCACAGAGGAGCCAGC 286
Db 413 GGGCAGCAGAACTGCCCGCCAGTGCAGACCCGCTTCAAGGCGCTCAAGGGGTGCGCGG 472
Qy 287 GATCCGGGGAGGAGGACGACGACACTGATGCCGATGATGTAGTACTTCAACTACC 346

2474 GGGATCTGCCCATCAATCTTTTCAGATCGTGTGAACCAAGTGTGCGGTGGGCTCTCGG 2533
Db AGGTTCCGACCTCTCAATCTTTTCGATCGTCTTACCAGGTTCTTCGGTGGGCTCTTGG 2686
QY TTCTGTTGAATCTTTTCAGCGCGGATTCGCCCTTTATGTTATGTTGCTACGGAGGGCGCT 2593
Db TTCAATTGAATTTCTTCAGCAACCACTGCCCTCTCTGTTATGGTATGGTGGACT 2746
QY CAAGTTCCTGGAGAGATTCTGTTATCAATCAACCACTATTTACCCTACTACCTCTCTCC 2653
Db AAGTTTCCTGGAAAGGTTTTCGTACATTAATCTCATCTGATACCTTTGGACATCTATCC 2806
QY GCTTCTAGTCTATTGATTTGCTGCTATCTCTCTGCTCACTGGAAAGTTCAATGCTC 2713
Db GCTCTTGGGCTATTGACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2866
QY AGAGATTAGCAACTTGGCCAGTATCTGTTTCAATTTGCTGCTTCTCTTCAATTTTCCAC 2773
Db AGAGCTTAACATTTTGGCAGCCTGTGTTTCAATTTGCTGCTTCTCTTCAATTTTCTAC 2926
QY TGATATCTTGAATGAGTGGAGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2833
Db GAGCATCTGGAATGAGTGGAGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2986
QY GTTCTGGGTCATTGGAGGCTGTCTTCAATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2893
Db GTTCTGGGTCATTGGAGGCTGTCTTCAATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 3046
QY GTGCTTGGGTCATTGGAGGCTGTCTTCAATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2953
Db GTGCTTGGGTCATTGGAGGCTGTCTTCAATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 3103
QY CGACTTTGTGAGTCTTACATTTTCAAGTGGAGCGAGCTTCTATCTCTCTCGAGCACCAT 3013
Db GAGATTCTCAGAGCTGTACACATTTTCAATTTGAGCGAGCCTTCTGATACCTTCGACACCT 3163
QY TTGATCATTAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3073
Db GTCTCTACTGAATTTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3223
QY CCAATCATGGGGGCGCTCTTTGGGAAGCTTCTTTTGGGCTTCTGCTGCTGCTGCTGCTGCT 3133
Db TGAATCATGGGGGCGCTCTTTGGGAAGCTTCTTTTGGGCTTCTGCTGCTGCTGCTGCTGCT 3283
QY ATACCATCTCTAAGGCTTATTTGGGAGGCAAAACCGACACCGGAGGATTTGCTGCTGCT 3193
Db TTACCCCTTCTCAAGGCTTCTGTTGGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 3343
QY CTGGGCTGCT 3253
Db CTGGGCTGCT 3403
QY TACCGCTCTGCTGCGCCCAATATCCAAACCTGTGGCATCACTGTGAGGAAATGGGAG 3313
Db TGCGAAGATGATGTTCCCTGTTGGAGGAGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTG 3463
QY TTTGTAGA 3321
Db CACGTGGA 3471

RESULT 8
AY103655
LOCUS Zea mays PCO096398 mRNA sequence.
DEFINITION Zea mays PCO096398 mRNA sequence.
ACCESSION AY103655
VERSION AY103655.1 GI:21206733
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3897)
Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 3897)
Coe, E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schubert, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
Location/Qualifiers
1..3897
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635496"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Query Match 35.0%; Score 1267.8; DB 11; Length 3897;
Best Local Similarity 63.9%; Pred. No. 6.7e-254; Indels 111; Gaps 8;
Matches 2108; Conservative 0; Mismatches 1082;
QY 47 GCTCTGCGCGCATGGACCGCGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 106
Db 203 GCTCTGCGCGCATGGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 262
QY 107 GGAAGTGTGCGAGATCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 166
Db 263 GCAGTGTGCGAGATCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 322
QY 167 CGCTGCGAGCTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 226
Db 323 GCGGTGCAACGAGTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 382
QY 227 GGGCACCAGCGCTGCTCCAGTGCAGAGCAAGTACAGCGCGCGCGCGCGCGCGCGCGCGCG 286
Db 383 GGGCACCAGCGAGTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 442
QY 287 GATCCGCGGGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 346
Db 443 TGTGACCGGTGACGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 502
QY 347 TGCATCTGGCACTGAGGACCGAGAGAGATGCTGACAGGATGCGAGCTGCGGCGAT 406
Db 503 CGGCCATGACTC-----GCAGTCTGTGGCGAGTCTTCTTCTTCTTCTTCTTCTTCTTCT 550
QY 407 GAACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATGGGCT 466
Db 551 GAGCTACGGCGGTGAGGTGA-----CCCTAATGGCGCG 584
QY 467 CTCGAAGTATGACAGTGGAGAGATCCCTTAGGGGATACGTCCTTCACTGACCAACAGCA 526
Db 585 CCACAAGCTTTCCAG-----CTCAACCCCAATGTTCCACTCTCTCAACAGGCGCA 634
QY 527 GATGTCAGGAGAAATCCCTGGAGCTTCGCTGATCATCATCATGATGTCCTTACCGGGA 586

Db 635 AATGGTGGATGACATCCACCGGAGCAGCACGGCTGCTTCTTTTCATGGGTGGTG 694
QY 587 CATCAGCAGAGTCTCCGTTTCCCTATGTGATCATTCACCAATCCGTCAGGGAGTT 646
Db 695 GGGAAAGAGGATACATCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCCAGGTC 754
QY 647 CTCGGCAGTATTTGGGAATTTCCCTGGAAAGAGAGAGTGTAGTGGCTGGAAATGAAGCA 706
Db 755 TATGGACCCATCCAAAGATCTTCTGCATATGGGTATGTAGTGTGCTTGGAGGAACG 814
QY 707 GGCACAGGGTGCATTCCTCATGATCATTAATGGGCAAGCATTTCTCCCTGAAAGTCCGGC 766
Db 815 GATGGGAATTTGGAAGCAGAGACAAGAGAGGATGCACAGACGGGAATGATGGTGGTG 874
QY 767 AGCTACTGACATCGATGCATCTACTGAATACACATGGAGAGCGTTTACTGATGATGA 826
Db 875 TGATGATGGTGAGATGCT-----GATCTACCACTAATGGATGA 913
QY 827 AACTGCCAGCCTCTATCTAGAAAAGTCCCATTTGCTTCTCCAAATAAATCCCTACAG 886
Db 914 AGCAAGACAACAACCTGCCAGGAAATTCACCTTCCATCAAGCCAGATTAAATCCATATAG 973
QY 887 AATGTCATTTGTCGGGTGGTGTCTTAAGCATCTTCCTGCACTACCGTCTCAAAA 946
Db 974 GATGATTAATCATTAATTCGGCTTTGGTGTGGGTTTCTTCCACTACCGAGTGATGA 1033
QY 947 TCCCTGTGCGTAATGCATACCACTGCTGCTTTTATCTGTTATATGAGATTGGTTTC 1006
Db 1034 TCCGGTGAATGATGATTTGCTTTTGGCGCATCTGTGTTATCTGTGAAATCTGGTTTC 1093
QY 1007 TTTATCTGGATCTGATGATTCCTCCGAAGTGGTTTCCAAATCAACCGGGACCTACCT 1066
Db 1094 CATGCTTGGATTCTTGATCAATTCCAAAGTGGTTCCCTATTTGAGAGAGACTTACCT 1153
QY 1067 TGATAGACTGGCTTTAAGTATGACCGGAGAGGTGAACCGTCTCAGTTGGCTGTGTA 1126
Db 1154 AGACCGGTGTCACTGAGGTTTCGACAAAGAGCCAGCCATCTCAACTTGCTCCAATTGA 1213
QY 1127 CATATTTGTCACTGACAGTCAGCCCTTGAAGAGGCCACTATCTGCTATGTATCTGATGA 1246
Db 1214 TTTCTTTGTAGTACGGTTGATCCCTTAAGGAACTCTCTTGGTCAACAATAATCTGT 1273
QY 1187 GCTATCCATTTCTGTGTGATTAATCCCGTGACAAAGGTCTCTTCTATGTATCTGATGA 1246
Db 1274 TCTATCTATCTCTTCCGTGGATTAATCCCTGTTGATTAAGTCTTCTGCTATGTTCTGATGA 1333
QY 1247 CGGAGCTTCAATGCTGACTTTGACGCAATGGCTGAGACTTCAGAGTTTGTCTAGAAATG 1306
Db 1334 TGGTGCTGCAATGCTTAACGTTTGAAGCATTAATCTGAAACATCTGAATTTGCAAGAAATG 1393
QY 1307 GGTACCATTTTGTGAAGAAGTATGACATTTGAACCCAGAGCTCCGAGTTTACTTTTGCCA 1366
Db 1394 GGTTCCTTTCTGAAAACGGTACATATTTGAACCTCGCGCTCCAGAGTGGTACTTCCAA 1453
QY 1367 GAAATTTGANTTACTGAAAGCAAAAGTCCAGCCTTCATTTGTTAAAGACCCCGGGCCAT 1426
Db 1454 GAAGATAGACTACTTGAAGACAAGGTGGCAGCAAACTTTCTTAGGGAGAGAGCAAT 1513
QY 1427 GAAGAGAGAAATGAGNAATTTAAATCAGGATAAATGCCCTAGTTTCTTAAGGCAATGAA 1486
Db 1514 GAAGAGAGAGTATGAGGAATTTCAAGGTGAGAAATCAATGCCCTTAGTTGCAAGCCAGAA 1573
QY 1487 AGTCCCGAGGAAGATGGATCATGCAAGATGGCACACCATGCGCCAGGAACAATACCAG 1546
Db 1574 AGTTCTTGAAGAAGGATGGACAATGCAAGATGGAAACCCCTCGGCTGGAAACAATGTTCC 1633
QY 1547 GGATCATCTCGGAATGATTCAGGTTTTCCTTGGTCAAGTGGTGGCCCTTGATCTAGGG 1606
Db 1634 TGATCATCTCGGAATGATTCAGGCTTCTTGGCCAAAGCGAGGCTTGTCTGTGAGGG 1693
QY 1607 TAATGAGCTCCCCGTTTAGTTTATGTTCTGCGTGAAGAGCGTCTCGGTTTCCAGCACCA 1666
Db 1694 AAATGAATGCCAGATTGGTTTATGTTTCTTAGAGAGAAACGACCGCTATTAACCATCA 1753

QY 1667 CAAGAAGGCTGTGGCCATGAATGCCCTTGTTCGTCTCAGCTGTCTTACTAATGACA 1726
Db 1754 TAAGAAGCTGTGTCTATGAATGCAATTCATTCGAGTCTGTCTGTACTAACAATGCTCC 1813
QY 1727 ATACATGTTGAATCTTGTATGTTGATCACTACATCAACAAGCAAGGCTGTCCGAGAAGC 1786
Db 1814 ATATTGTTTAAACTTGGATTTGATCACTACATCAACAAGCAAGGCTATAAAGGAAGC 1873
QY 1787 TATGTCCTTCCCTAATGATCCAAACCTAGTCCGCAAGTCTGTATGTGTCAGTCTCCACA 1846
Db 1874 AATGTCCTTTATGATGAGCCCTTTACTAGAAAGAGTTTGTCTATGTATCAGTCTCCCTCA 1933
QY 1847 AAGTTTGTATGGGATTTAGGAATGATCGATATGCAAAACAGCAACACTGTCTTTTTTGA 1906
Db 1934 AAGATTGATGGATTTGATCGCCATGACCGATATGCTAACCCGAATGTTGTCTTTTTTGA 1993
QY 1907 TATTAACCTTGGGGCTTTGACGGCATTCAGGCAATCAAGCAAGTTTATGTGGGAACCTGTTGTGT 1966
Db 1994 TATCAACATGAAGGTTTGGATGGTATTCAGGGTCCAAATTTATGTTGGTACTGGATGTGT 2053
QY 1967 TTTCAACAGAACAGCTATCTATGGTTTATGAGCCCCCAATTAAGGCGAAGAGCCA----- 2021
Db 2054 ATTTAGAGGCGAGGCATTAATATGTTATGATGCCCCCAAAACAAGAGGCCACCATCAAG 2113
QY 2022 -----GGTTTCTGGCATCACTATGTGGGGGCAAGAAAGGC 2059
Db 2114 GACTTGCAACTGCTGGCCCCAAGTGTGCTTTTGTGCTGCTTTGGCAATAGGAAGCA 2173
QY 2060 AAGCAAGTCAAGAAAGAGGCTCAGATTAAGAAAAG-----TCGAACAAGCATGTGA 2113
Db 2174 AAAGAAGACTACCAAAACCAAAACAGAGAAGAAAAGTTATTTTTCAGAAAGAAAGA 2233
QY 2114 CAGTTCTGTCTTCAGTATTCAATCTCGAAGACATAGAGAGGCTGTGAAGGTGCTGGGTT 2173
Db 2234 GAACCAATCCCTGCGATGCTCTTGGTGAATTTGACGAAGCTGCTCCAGGAGC----- 2287
QY 2174 TGATGATGAGAAATCAGTTCTCATGTCTCAATGAGCTTAGAGAAAGATTGGCCAGTC 2233
Db 2288 TGAGAAATGAAAGGCGGTATTTGTAATCAACAATAATAGAAAGAAATTTGGCCAAATC 2347
QY 2234 AGCAGATTTGTGCTCCACTCTGATGGAATATGGTGTGTTCTCTCAGTCTCCACTCC 2293
Db 2348 TTTCTGTTTTGTATCATCCACACTCTCGAAGTGGTGAACCTTTGAAGAGTGCAGTCC 2407
QY 2294 AGAATCTCTTTGAAAGAAAGCTATCCATGTCATAGTTGTGGCTATAGGCAAGTCTCA 2353
Db 2408 TGTCTCTCTTTGAAAGAGAGCTATACATGTCATGTTGTGTATGAGACAGACAGA 2467
QY 2354 ATGGGGAACCTGAGATTTGGATCTATGGATCTGTCTACAGAAGATATTTCTTACTGGATT 2413
Db 2468 CTGGGAAAAGAGATTGGCTGGATCTATGGATCAGTTACAGAAGATATTTCTAAGTGT 2527
QY 2414 CAAGATGACGCAAGAGGCTGGCGTTGAGTCTATGATGCCAAGCGCCAGCCAGCTTTCAA 2473
Db 2528 CAAGATGATTTGTATGTTGGCGGTCAATTTACTGCTATCTAAAACGGGTGTCATTCAA 2587
QY 2474 GGGATCTGCCCCCAATCAATCTTTTCAGATCGTCTCAACCAAGTGTGGCTGGGCTCTCGG 2533
Db 2588 AGGTTCTGCACTCTGAATCTTTTCAGATCGTCTTCAAGGTGCTTCGGTGGGCTCTGG 2647
QY 2534 TTTCTGTTGAAATTTCTTTTCAGCGCGCAATTCCTTATGGTATGGCTACGAGGCGCGCT 2593
Db 2648 GTCTATTGAGATCTTCTTCAGCAATCATTTGCCCTCTTTGGTATGGGTATGTTGGCGGTCT 2707
QY 2594 CAGTTCTCGAGAGATTTCCTTACATCAACACCAATTTACCCACTAATCTCTCTCC 2653
Db 2708 GAAATTTTGGAAAGATTTTCTTACATCAACTCCATCGTGTATCTTGGACATCTATTCC 2767
QY 2654 GCTTCTAGTCTATTGTATTTGCTGCTATCTGTCTCACTGGAAGTTTCATCATGCC 2713
Db 2768 CCTCTGGCTTACTGTACATTGCTGCCATCTGTTTATTGACAGGGAAATTTTATCACTCC 2827

2714 AGAGATTAGCACTTGGCCAGTATCTGGTTCAATGGCTCTTCTTCAATTTTCGCCAC 2773
|||||
2828 AGAGCTGAATAATGTTCACAGCTGTGGTTCATGTACATTTTATCTGCATTTTGGCTAC 2887
|||||
2774 TGGTATCCTTTGAGATGAGGTGGTGTGGCAATGACAGTGGTGGAGGAATGAACA 2833
|||||
2888 GAGCATCTAGAAATGAGATGAGTGTGGATTTGATGACTGGTGGAGGAATGAGCA 2947
|||||
2834 GTTCTGGGTCATTTGGAGGTATCTGCAATCTGTTTGGCTTTTCAGGGCTTCTGAA 2893
|||||
2948 GTTCTGGGTCATTTGGAGGTGTGCTCTCACACCTTTTGTGTGTTCCAGGGACTTCTCA 3007
|||||
2894 GGTCTTGGCGGTATCGACCAACTTCACTGCTCACTCAAGGCTCAATGACGAAGG 2953
|||||
3008 GGTCTAGTGTGGTATGATACAGCTTCCCGGTGACATCAAGG---GTGGAGATGATGA 3064
|||||
2954 CGACTTTGTGAGCTCTACATGTTCAAGTGGAGCGCTTCTCATCCCTCCGACGACCAT 3013
|||||
3065 GAGTTTCTAGAGCTATATACATTCATGAGTACCTTATGATACCTTCTTACCACCTT 3124
|||||
3014 TTGATCATTAACATGGTGGTGTGCTGCGACCTCTAGCCATCAAGAGTGGTTA 3073
|||||
3125 GCTTCTATTGAATCTTCAATGCTGTGCTGCGCTTTCAAATGCCATCAATAACGATA 3184
|||||
3074 CCAATCATGGGCGCTCTTTGGAGGCTCTTCTTGGCTTCTGCGTGTGATTTGCTCACTT 3133
|||||
3185 TGAGTCATGGGCGCTCTTTGGAGGCTCTTCTTGGCTTCTGCGTGTGATTTGCTCATCT 3244
|||||
3134 ATACCATTCCTCAAGGCTTATGGGCGAGCAAAACCGCACCGACGATGTGTCATCGT 3193
|||||
3245 TTATCCCTTCTCAAGGTTTGGTGGAGGCAAAACGAGCACCAACGATTTGTCATCGT 3304
|||||
3194 CTGGGTGCTCTCTGCTCTATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3253
|||||
3305 CTGCTCATTTCTGCTGCTTCAATCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCT 3364
|||||
3254 TACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3313
|||||
3365 TGGAGGATGATGCTCGCTTCTTGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3424
|||||
3314 T 3314
3425 T 3425

CD726831 1874 bp mRNA linear EST 26-JUN-2003
EST027 Cucurbita pepo testa subtracted cDNA Cucurbita pepo cDNA
Clone CES similar to Cellulose synthase, mRNA sequence.
CD726831
CD726831.1 GI:32277678
EST.
Cucurbita pepo
Cucurbita pepo
Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
1 (bases 1 to 1874)
Bezold,T.N., Mathews,D., Loy,J.B. and Minocha,S.C.
Molecular analysis of the hull-less seed trait in pumpkin:
Expression profiles of cell wall related genes during development
Unpublished (2003)
Contact: Subhash Minocha, Dennis Mathews, Brent Loy, Todd Bezold
Dr. Minocha
University of New Hampshire
Rudman Hall, Durham, NH 03824, USA
Tel: 603 862 3840
Fax: 603 862 3784
Email: sminocha@cisunix.unh.edu
Degenerate primers and Tag were used to amplify cDNA for TOPO TA
(Invitrogen, Carlsbad, CA) cloning. Sequencing was performed three
times using the Dyanamic ET Terminator Sequencing kit (Amersham

Pharmacia Biotech Inc, Piscataway, NJ). Sequences were identified
by NCBI BLAST(X).
Location/Qualifiers
1. 1874
/organism="Cucurbita pepo"
/mol_type="mRNA"
/db_xref="taxon:3663"
/clone="CES"
/dev_stage="20 days post-anthesis"
/clone_lib="Cucurbita pepo testa subtracted cDNA"
/note="Organ: Testa; Total RNA was isolated from 20 day
post-anthesis testa tissue and used in a subtraction
hybridization procedure as according to the Clontech
PCR-Select cDNA Subtraction kit (PT1117-1) (Clontech, Palo
Alto, CA)."

ORIGIN

Query Match 24.3%; Score 882; DB 14; Length 1874;
Best Local Similarity 67.7%; Pred. No. 2.4e-173; Indels 6; Gaps 2;
Matches 1268; Conservative 0; Mismatches 600;

QY 1016 GATACGTGATCAGTCCCGAAGTGGTTTCCAAATCAACCGGAGACCTTACCTTGATAGACT 1075
Db 1 GGTGCTGGACCATTCCTTAATGGTTACTGTCCTGAGGACACATTTATTGACAGATT 60

QY 1076 GGTCTTAAAGTATGACCGGAGAGGTGAACCGTCTCAGTTGGCTGCTGTTGACATATTGT 1135
Db 61 ATCAGCCAGGTTTGAAGAGAGAGGAGAACCTTCCAGCTTGTGCTGTGGATTTTGT 120

QY 1136 CAGTACAGTCGACCCCTTGAGGAGGACCATCTATGTCACGCCAACACTGTGCTATCCAT 1195
Db 121 GAGTACCGTTGATCCGTTGAAGAACCTCTCTAAATCACTGCGAATACCGTGTTCAT 180

QY 1196 TCTTGTCTTGTGATATCCCGTGGACAAAGTCTCTTGTCTATGTATCTGATGACGAGGCTC 1255
Db 181 CCTTGTGTGAGTATCTCTGTAAGTCTCTCTGTAATGTGTCGATGACGCTGAGC 240

QY 1256 AATGCTGACTTTTGAACCATTTGGCTGAGATCTCAGTTTGTCTAGGAAATGGGTACCAT 1315
Db 241 TATGCTCACATTTGAATCTCTAGTTGAAACAGCTGACTTTTGCAGGAAATGGGTTCAT 300

QY 1316 TGTGAAGATGACATTTGAACCCAGAGCTCCGAGTTTCTTCTTCCAGAAATTTGA 1375
Db 301 CTGCAAAAAATTTCCATTTGAACCTCGAGTCTTCTGATTTTCTTCTCCAGAAATAGA 360

QY 1376 TTACTGAAAGACAAAGTCCAGCTTTCATTTTGTAAAGACCGCGGCCCATGAGAGAGA 1435
Db 361 CTATTGAAAGATGAAATGACATCTTCTTTGTAAAGGAACTGATAGAGCTATGAGAGAGC 420

QY 1436 ATATGAAGAAATTTAAATCAGGATAAATGCCCTAGTTTCTTAAAGCATTTGAAGTCCCGA 1495
Db 421 TTATGAAGAGTTTAAAGTACGAGTCAATGCTTTTGGTGGCAAGGCGCAAAAAACACCTGA 480

QY 1496 GGAAGGATGATCATGCAAGATGACACCATGCGCAGGAAACAAATACAGGGATCATCC 1555
Db 481 TGAAGGCTGGTCCATGCAAGATGAAACAGCTTGGCCAGGAAACAAATCCACGTGATCACC 540

QY 1556 TGGAAATGATTCAGGTTTCTTGTGTCAGTGGTGGCTTGTATCTAGGGTATGAGCT 1615
Db 541 TGGATGATTCAGGTTTCTTGTGAAATACCGCGGCCCATGACGCTAGAGGGAATGAACT 600

QY 1616 CCCCGTTTATGTTGCTCTGCTGAAAGCGTCTGGTTCAGGACCAACAGAGG 1675
Db 601 TCCTCGACTAGTCTCTCCAGAGAGAGAGGCTGTTTACCAGCATCACAAAAAGC 660

QY 1676 TGGTGCATGATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1735
Db 661 TGGTGTGAAATGCTCTGGTAAGATATCTGAGTCTCTCACAATGCCCTTCTTCTTCT 720

QY 1736 GAATCTGATTTGATCATCTACATCAACAAAGAGGCTGTCCAGAGCTATGTGCTT 1795
Db 721 TAACTTCTGCTGATCATCTATGTTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 780

FEATURES
source

```
QY 1796 CTTAATGGATCCAAACCTAGTCCGCAAGTCTGTTATGTGAGTTCCTCCCAAGGTTTGA 1855
Db      |||||
QY 1856 TGGGATTTGATAGGATCATGATATGCAACAGAGAACACATGCTCTTTTGTATTAACATT 1915
Db      |||||
QY 1916 GAGGGGCTTTCAGGGCATTCAGGACACAGTTTATGTGGGAACATGCTGTTTTCACAG 1975
Db      |||||
QY 1975 GAAAGGATTTGGATGGCATTCAGGACACAGTTTATGTTGGTACAGGTTGCTTTTCAATAG 960
QY 1976 AACAGTATCTAGTTATGAGTCCCAATTAAGGCGAAGAACCC---AGGTTCTTGGC 2032
Db      |||||
QY 2032 ATCACTATGTGGGGGCAAGAGGCAAGCAAGTCAAGAAAGAGGCTCAGATAAGAA 2092
Db      |||||
QY 2092 AATCTTCTCTGTGTGGCTGTCTTGTCTGTCTGTCTTCCCAAGAGATCTCAAAAGA 1080
QY 2093 AAAGTCGAACAAGCATGTGACAGTTCTGTCTCAGTATTCAACTCGAAGACATAGGA 2152
Db      |||||
QY 1081 TCCGACTGAGATTCAGAGAGATCCAAAAGAGAGAGCTTGATGCTGCAATCTTTAACCT 1140
QY 2153 GGGTGTGAAGTGTCTGGGTTTGTATGATGAGAAATCAGTTCTCATGTCTCAATAGAGCTT 2212
Db      |||||
QY 1141 CAGGGAATAGATAATATGATGAGTATGAGAGATCAATCTGATTTCTCACTGAGCTT 1200
QY 2213 AGAAGAGATTTGGCAGTACAGCATTTGTTGCTCCTCAGTCTGATGGAATATGCTG 2272
Db      |||||
QY 1201 TGAGAAACTTTTCGGATTCGTCTGTGTTATTCAGTCTAGCTAATGGAATGGCGG 1260
QY 2273 TGTTCCTCAGTCTCCTCAGTCCAGAACTCTTTTGAAGAGCTATCCATGTCATAGTTG 2332
Db      |||||
QY 1261 AGTTCCGGAATCTGAAATCCCTCGACTTGATCAAGGAAGCAATTCATGTCATTAGCTG 1320
QY 2333 TGGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTTGTTGATCTATGGATCTGTCTAC 2392
Db      |||||
QY 1321 TGGTTATGAAGAGAAGTCCACTTGGGAAAAGAGATTTGTTGGATATATGGGTGAGTGAC 1380
QY 2393 AGAAGATATCTTACTGGATTCAGAGTACGAGACGAGGCTCGGCTCAGTCTATTTGAT 2452
Db      |||||
QY 1381 TGAGATATCTTAAACGGGTTTCAAGATGCAITTCGAGGGGTGAGGTCTTACTGCTAT 1440
QY 2453 GCCCAAGCGCCAGCTTTCAAGGGATCTGCCCCCATCAATCTTTCAGATCGTCTGAACA 2512
Db      |||||
QY 1441 GCCATTTAGGCAGCATTCAGAGGTTCGCCAATTAACCTTTCTGATCGTCTCCACCA 1500
QY 2513 AGTGTGCGGTGGGCTCTGGTCTCTGTTGAAATCTTTTCAGCCGGCATTTGCCCTTATG 2572
Db      |||||
QY 1501 AGTTCTTCGATGGGCACCTTGGATCTGTTGAGATTTTCTTAGCACACACTGTCCATATG 1560
QY 2573 GTATGGCT---ACGAGAGGGGCTCAAGTTCCTGGAGAGATTCGTTTACATCAACACCAC 2629
Db      |||||
QY 1561 GTATGGAATTTGAGGCGCGCGCTCAAAATGGCTCCAAAGAAATGGCTTACATAAACCACAT 1620
QY 2630 CATTTACCCACTAACTCTCTCCGCTCTAGTCTATTTGATATGTCCTGCTATCTGCT 2689
Db      |||||
QY 1621 TGTCTATCCCTTCACTCGCTCCCTCTCTGTTTACTGCTCATGCTGCAATCTGCT 1680
QY 2690 GCTCATGGAAGTTTCAATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGTTTCATGTC 2749
Db      |||||
QY 1681 GCTCACAGGAAAGTTTCAATTCACACGCTCTCGAACCTAGCAAGTACCCCTTTTCTCGG 1740
QY 2750 GCTCTTCTTTCAATTTTCGGCATGTTGATCTGAGATGAGGTGGATGGTGTGGCAT 2809
Db      |||||
QY 1741 TCTCTTCTTGTCCATCAATTCACGAGTGTCTCGAGCTGCTGGAGTGGTGTAGCAT 1800
QY 2810 TGACAGTGGTGGAGGATGACAGTCTCTGGGTCAATGGAGGTATCTCTGACATCTGTT 2869
Db      |||||
QY 1801 CGAAGATATATGGCGTACAGAGCAATCTGGGTAAATCGGAGGCGTCCCGGCACATCTCTT 1860
QY 2870 TGCGGTCTTTCCAGG 2883
```

Db 1861 TGCGGTCTTCCAAG 1874

RESULT 10
BM816138
LOCUS
DEFINITION

BM816138 880 bp mRNA linear EST 05-MAR-2002
HC109B12_T3.ab1 HC Hordeum vulgare subsp. vulgare cDNA clone
mayb), (AF200529) cellulose synthase-5 [Zea mays], (AF200533)
cellulose synthase-9 [Zea mays], cellulose synthase catalytic
subunit [Arabidopsis thaliana], unnamed protein product [Arabidopsis
thaliana], mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BM816138.1 GI:19152152
EST.
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.

REFERENCE
AUTHORS

Ozturk,N.Z., Michalowski,C.B., Brazille,S., Borchert,C.,
Palacio,C., Normand,C., Murphy,C., Kelley,R., Sant,S.A.,
McLaughlin,H., Fredrickson,M.A. and Bohnert,H.J.
Monitoring large-scale changes in transcript abundance in drought-
and salt-stressed barley
Unpublished (2002)
Contact: Mark A. Fredrickson
Plant Biology
University of Illinois
1201 W Gregory Dr, Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.

TITLE

JOURNAL

COMMENT

FEATURES

Source

1..880
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/strain="cv tokak"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HC109B12_T3.ab1"
/tissue_type="Root"
/dev_stage="3 week old"
/clone_lib="HC"
/note="6 and 10 hour drought stress by placing plants on
moist paper (75% rel. humidity) in light"

ORIGIN

Query Match 21.0%; Score 762.2; DB 12; Length 880;
Best Local Similarity 96.2%; Pred. No. 2.1e-148;
Matches 801; Conservative 0; Mismatches 30; Indels 2; Gaps 2;

QY 2469 TTCAGGGATCTGCCCCCATCAATCTTCAGATCGTCTGAACCAAGTCTGCGTGGGCT 2528
Db |||||

QY 2529 CTCGGTCTGTGTAATTTCTTTTCAGCGGCATTTGCCCTTATGTTATGCTACGGAGG 2588
Db |||||

QY 107 CTTCGGTCCGTGTAATTTCTTTTCAGCGGCATTTGCCCTTATGTTATGCTACGGAGG 166
Db |||||

QY 2589 CGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACCATTTACCCCAACCTCT 2648
Db |||||

QY 167 CGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACCATTTACCCCAACCTCT 226
Db |||||

QY 2649 CTCGGCTTCTAGTCTATTTATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2708
Db |||||

QY 227 CTCGGCTTCTAGTCTATTTATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 286
Db |||||

QY 2709 ATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTCATTTGGCTCTTCCCTTTCAATTTTC 2768
Db |||||

QY 287 ATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTCATTTGGCTCTTCCCTTTCAATTTTC 346
Db |||||


```

RESULT 12
BQ802778      762 bp      mRNA      linear      EST 30-JUL-2002
LOCUS        WHE2829_H06_O11Zs Triticum monococcum vernalized apex cDNA library
DEFINITION   Triticum monococcum cDNA clone WHE2829_H06_O11, mRNA sequence.
ACCESSION   BQ802778
VERSION     BQ802778.1 GI:22017747
KEYWORDS     EST.
SOURCE      Triticum monococcum
ORGANISM    Triticum monococcum
REFERENCE   1 (bases 1 to 762)
AUTHORS     Anderson, O.D., Chao, S., Crossman, C., Dubcovsky, J., Echenique, V.,
            Lazo, G.R., Pham, J., Rausch, C.J., Stamova, B., Wilson, C. and Woo, J.
TITLE       The structure and function of the expressed portion of the wheat
            genomes - Vernalized apex cDNA library from Triticum monococcum
JOURNAL     Unpublished (2002)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oanderson@pw.usda.gov
            Sequences have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: SK primer.
FEATURES             Location/Qualifiers
                     1..762
                        /organism="Triticum monococcum"
                        /mol_type="mRNA"
                        /cultivar="G3116"
                        /db_xref="taxon:4568"
                        /clone="WHE2829_H06_O11"
                        /tissue_type="Vernalized apex"
                        /dev_stage="One month old plants"
                        /lab_host="E. coli XL0LR"
                        /clone_lib="Triticum monococcum vernalized apex cDNA
                        library"
                        /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                        Site 1: EcoRI; Site 2: XhoI; One-month old plants were
                        subjected to vernalization treatment by placing them in
                        the cold room at 6 C, under 15hr light/9hr dark condition.
                        Total RNA was prepared from apex tissue extracted from
                        plants with no cold treatment; and from plants with
                        2-week, 4-week and 6-week cold treatment separately. Equal
                        amount of total RNA was pooled from all four samples, a
                        cDNA library was made using pooled polyA RNA and cDNA
                        clones were in vivo excised at the University of
                        California, Davis (V. Echenique, B. Stamova, J.
                        Dubcovsky). Plasmid DNA preparations and DNA sequencing
                        were performed in the OD Anderson lab (all other
                        authors)."
```

ORIGIN

```

Query Match      20.7%; Score 750.4; DB 13; Length 762;
Best Local Similarity 99.9%; Pred. No. 6e-146;
Matches 751; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 625 CACCAATCCGTCAGGAGTTCTCCGGCAGTATTGGGAATGTCCTGGAAGAGAGAG 684
Db 11 CAGCAATCCGTCAGGAGTTCTCCGGCAGTATTGGGAATGTCCTGGAAGAGAGAG 70
Qy 685 TTGATGGCTGGAATGAAGCAGACACAGGGTGGGATTCCTCCATGACTAATGGCAAGCA 744
Db 71 TTGATGGCTGGAATGAAGCAGACACAGGGTGGGATTCCTCCATGACTAATGGCAAGCA 130
Qy 745 TTGCTCCCTCTGAAGTTCGGCAGCTACTGACATGATGATGATGATGATGATGATGATG 804
Db 131 TTGCTCCCTCTGAAGTTCGGCAGCTACTGACATGATGATGATGATGATGATGATGATG 190
```

```

Qy 805 AAGACGCTTTACTGAATGATGAACCTCGCCAGCCTCTATCTAGAAAGTCCCATTCGTT 864
Db 191 AAGACGCTTTACTGAATGATGAACCTCGCCAGCCTCTATCTAGAAAGTCCCATTCGTT 250
Qy 865 CTCCAAAATAAATCCCTACAGAAATGGTTCATGTTCTTCGGGTGGTGTCTTAAGCATCT 924
Db 251 CTCCAAAATAAATCCCTACAGAAATGGTTCATGTTCTTCGGGTGGTGTCTTAAGCATCT 310
Qy 925 TCTGCACTACCGTCTCACAATAATCCCTGCTGCTAAATGCATACCCACATGTCGCTTTTATCTG 984
Db 311 TCTGCACTACCGTCTCACAATAATCCCTGCTGCTAAATGCATACCCACATGTCGCTTTTATCTG 370
Qy 985 TTATATGTGAGATTTGGTTTGGCTTTTATCTGGAATCTGGATCAGTTCCTCCGAAGTGGTTTC 1044
Db 371 TTATATGTGAGATTTGGTTTGGCTTTTATCTGGAATCTGGATCAGTTCCTCCGAAGTGGTTTC 430
Qy 1045 CAATCAACGGGAGACCTACCTTGATAGACTGGCTTTTAAGGTATGACCGAGAGGTGAAC 1104
Db 431 CAATCAACGGGAGACCTACCTTGATAGACTGGCTTTTAAGGTATGACCGAGAGGTGAAC 490
Qy 1105 CCTCTCAGTTGGTCTGCTGTGACATATTTGTGACAGTACAGTCCGACCCCTTGAAGAGCCAC 1164
Db 491 CCTCTCAGTTGGTCTGCTGTGACATATTTGTGACAGTACAGTCCGACCCCTTGAAGAGCCAC 550
Qy 1165 CTATCGTCACTCCCAACACTGTCGTATCCATTTCTTGTGTTGATTATCCCGTGGACAAGG 1224
Db 551 CTATCGTCACTCCCAACACTGTCGTATCCATTTCTTGTGTTGATTATCCCGTGGACAAGG 610
Qy 1225 TCTCTTGTGATGATCTGATGACGGAGCTTCAATGCTGACTTTTGACGATTTGGCTGAGA 1284
Db 611 TCTCTTGTGATGATCTGATGACGGAGCTTCAATGCTGACTTTTGACGATTTGGCTGAGA 670
Qy 1285 CTTACAGATTTGCTAGGAAATGGGTACCATTTGTGAAAGATGATGACATTGAACCCAGAG 1344
Db 671 CTTACAGATTTGCTAGGAAATGGGTACCATTTGTGAAAGATGATGACATTGAACCCAGAG 730
Qy 1345 CTCCTGAGTTTACCTTTTGGCCAGAAAATGTAT 1376
Db 731 CTCCTGAGTTTACCTTTTGGCCAGAAAATGTAT 762

RESULT 13
BQ368813      870 bp      mRNA      linear      EST 22-OCT-2001
LOCUS        HVSME10020M13f Hordeum vulgare 20 DAP spike EST library HVCNDA0010
DEFINITION   (20 DAP) Hordeum vulgare subsp. vulgare cDNA clone HVSME10020M13f,
            mRNA sequence.
ACCESSION   BQ368813
VERSION     BQ368813.1 GI:13257914
KEYWORDS     EST.
SOURCE      Hordeum vulgare subsp. vulgare
ORGANISM    Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
            1 (bases 1 to 870)
REFERENCE   Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D.,
            Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W.,
            Fenton, R.D., Close, S.J., Oates, R. and Main, D.
            Development of a genetically and physically anchored EST resource
            for barley genomics: Morex 20 DAP spike cDNA library
            Unpublished (2001)
            Contact: Wing RA
            Clemson University Genomics Institute
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hg bases = 537
            Seq primer: AATTAACTCCTCACTAAAGG
            High quality sequence stop: 777.
```

FEATURES	
Location/Qualifiers	Source
<p>1..870 /organism="Hordeum vulgare subsp. vulgare" /mol_type="mRNA" /cultivar="Morex" /sub_species="vulgare" /db_xref="taxon:112509" /clone="HVSMEI0020M13f" /tissue_type="20 DAP spike" /lab_host="SOLR" /clone_lib="Hordeum vulgare 20 DAP spike EST library HVCDNA0010 (20 DAP)" /notes="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI; Plants were grown in the greenhouse at the University of California, Riverside (Fenton, SJ Close, TJ Close). Whole spikes with awns trimmed were collected at 20 DAP (Fenton). Total RNA was prepared, poly(A) RNA was purified, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids in the TJ Close lab at the University of California, Riverside (Choi). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begun, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see http://www.genome.clemson.edu/projects/barley/. To order this clone see http://www.genome.clemson.edu/orders Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/ggpages/bgn/31/clover.html)"</p>	
ORIGIN	
Query Match	20.1%; Score 729.4; DB 12; Length 870;
Best Local Similarity	93.8%; Pred. No. 1.5e-141;
Matches 813; Conservative	0; Mismatches 48; Indels 6; Gaps 5;
QY	1747 GTGATCTACTACATCAACAAGCGTGTCCGAGAAGCTATGTGCTTCCTTAATGGATC 1806
DB	1 GTGATCTACTACATCAACAAGCGTGTCCGAGAAGCTATGTGCTTCCTTAATGGATC 60
QY	1807 CAACCTAGTCCCAAGTCTGTATGTGCAGTCCCACAAAGTTTGATGGATTGATA 1866
DB	61 CAACCTAGTCCCAAGTCTGTATGTGCAGTCCCACAAAGTTTGATGGATTGATA 120
QY	1867 GGATGATCATATGAACAAGCAAGCACTGTCTTTTTTGATATAAATCTCAGGGGCCITG 1926
DB	121 GGATGATCATATGAACAAGCAAGCACTGTCTTTTTTGATATAAATCTCAGGGGCCITG 180
QY	1927 ACGGCATTCAGGACCAGTGTATGTGGAACTGGTTGTGTTTTCACAGAACGTATCT 1986
DB	181 ATGGCATTCAGGGCCAGTTTATGTGGAACTGGTTGTGTTTTCACAGAACGGCTATCT 240
QY	1987 ATGGTTATGACCCCCATTAAGCGCAGAGCCAGGTTTCTGGCATCACTATGTGGGG 2046
DB	241 ATGGTTATGACCCCCCATTAAGCGCAGAGCCAGGTTTCTGGCATCACTATGTGGGG 300
QY	2047 GCAAGAAGCGGCAAGCAAGTCAAAAAGGAGGCTCAGATAAGAAAAAGTGCACCAAGC 2106
DB	301 GCAAGAAGCGGCAAGCAAGTCAAAAAGGAGGCTCAGATAAGAAAAAGTCAAACAAGC 360
QY	2107 ATGTGGACAGTCTCTCCAGTATTCATCTCGAAGCATAGAGGGGGTTTGAAGGTG 2166
DB	361 ATGTGCAGAGTCTCTCCAGTATTCATCTCGAAGCATAGAGGGGGTTTGAAGGTG 420
QY	2167 CTGGGTTTGATGAGAAATCAGTTCTCATGTCTCAATGAGCTTAGAAGACAGATTTC 2226
DB	421 CTGGGTTTGATGAGAAATCAGTTCTCATGTCTCAATGAGCTTAGAAGACAGATTTC 480

```
QY 2974 TGTTCAGTGGACGACGCTCTTCATCCCTCCGACGACCAATTTTGATCAATTAACATGGTTG 3033
Db 653 TGTTCAGTGGACGACGCTCTTCATCCCTCCGACGACCAATTTTGATCAATTAACATGGTTG 594
QY 3034 GTGTGCTGTGCTGGACGCTCTTAGCCATCAACAGTGGTTACCAATCATGGGGCCGCTCT 3093
Db 593 GTGTGCTGTGCTGGACGCTCTTAGCCATCAACAGTGGTTACCAATCATGGGGCCGCTCT 534
QY 3094 TTGGGAAGCTCTCTTTGCTTCTGGGTGATTTGCTCACTATATACCAATCTCTCAAGGTC 3153
Db 533 TTGGGAAGCTCTCTTTGCTTCTGGGTGATTTGCTCACTATATACCAATCTCTCAAGGTC 474
QY 3154 TTATGGGAGGCAAAACCCACACACGAGATTGTCTGCTGGGTGCTCTCTCGCTT 3213
Db 473 TTATGGGAGGCAAAACCCACACACGAGATTGTCTGCTGGGTGCTCTCTCGCTT 414
QY 3214 CTATCTTCTCTTCTGCTGGGTGCTGGTGTGATTCATTCATCCCTCTCGTGGCCAA 3273
Db 413 CTATCTTCTCTTCTGCTGGGTGCTGGTGTGATTCATTCATCCCTCTCGTGGCCAA 354
QY 3274 ATATCCAAACCTGTGGCATCAACTGCTAGGAAAGTGGGAGTTGTAGACAGAAATAT 3333
Db 353 ATATCCAAACCTGTGGCATCAACTGCTAGGAAAGTGGGAGTTGTAGACAGAAATAT 294
QY 3334 AACAGTATCGAGCAACAAACCCGCGAGCCAGAGAAATTTATTTGTTGGGTTGTGAATTA 3393
Db 293 AACAGTATCGAGCAACAAACCCGCGAGCCAGAGAAATTTATTTGTTGGGTTGTGAATTA 234
QY 3394 CTACGTTTGGAGAACTTGTCAAATTTGAGAAACACATTTGTAATAGATGAATAGACC 3453
Db 233 CTACGTTTGGAGAACTTGTCAAATTTGAGAAACACATTTGTAATAGATGAATAGACC 174
QY 3454 ATCTACCGTTTTCATGAGTAAAGCTCTCTTTTGTGAAACAAAGGAATCTCATTTGTA 3513
Db 173 ATCTACCGTTTTCATGAGTAAAGCTCTCTTTTGTGAAACAAAGGAATCTCATTTGTA 114
QY 3514 AACCTATAGGAATTTTCTATGAGCACTTTGGAATTTAGGAATGAGCACTATGAATGTT 3573
Db 113 AACCTATAGGAATTTTCTATGAGCACTTTGGAATTTAGGAATGAGCACTATGAATGTT 54
QY 3574 GTATTTATTTATATATAAATTTCCCTGCTTCCACATTTTGGAGGATTTT 3626
Db 53 GTATTTATTTATATATAAATTTCCCTGCTTCCACATTTTGGAGGATTTT 1

RESULT 15
CD930401 701 bp mRNA linear EST 15-JUL-2003
LOCUS GR45.111C19F010418 GR45 Triticum aestivum cDNA clone GR45111C19,
DEFINITION mRNA sequence.
ACCESSION CD930401
VERSION CD930401.1 GI:32778165
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 701)
Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).
FEATURES
Location/Qualifiers
1..701
/organism="Triticum aestivum"
```

```
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="GR45111C19"
/tissue_type="grain (45 degrees per day after
pollination)"
/clone_lib="GR45"
```

ORIGIN

```
Query Match 19.0%; Score 689.8; DB 14; Length 701;
Best Local Similarity 99.0%; Pred. No. 2.7e-133;
Matches 694; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2324 CATAGTTTGTGGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTGGTTGATCTATGG 2383
Db 1 CATAGTTTGTGGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTGGTTGATCTATGG 60
QY 2384 ATCTGTCAAGAGATATTTCTTACTGGATTCAAGATGCAAGGCTGGGTTTCAGT 2443
Db 61 ATCTGTCAAGAGATATTTCTTAACTGGATTCAAGATGCAAGGCTGGGTTTCAGT 120
QY 2444 CTATTGCATGCCAACGCGCCAGCTTTCAAGGGATCTGCCCCATCAATCTTTTCAGATCG 2503
Db 121 CTATTGCATGCCAACGCGCCAGCTTTCAAGGGATCTGCCCCATCAATCTTTTCAGATCG 180
QY 2504 TCTGAACCAAGTGTGCGGTGGGCTCTCGGTTCTCTGTTGAAATCTTTTCAGCCGCAATTG 2563
Db 181 TCTGAATCAAGTGTGCGGTGGGCTCTCGGTTCTCTGTTGAAATCTTTTCAGCCGCAATTG 240
QY 2564 CCCCTTATGGTATGCTACGAGGGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAA 2623
Db 241 CCCCTTATGGTATGCTACGAGGGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAA 300
QY 2624 CACCAACCATTTACCACATAACCTCTCTCCCGCTTCTAGTCTATTATATTCCTGCTAT 2683
Db 301 CACCAACCATTTACCACATAACCTCTCTCCCGCTTCTAGTCTATTATTCCTGCTAT 360
QY 2684 CTGCTGCTCACTGAAAGTTTCATGCTCCAGAGATTAGCAACTTTGGCCAGTATCTGGTT 2743
Db 361 CTGCTGCTCACTGAAAGTTTCATGCTCCAGAGATTAGCAACTTTGGCCAGTATCTGGTT 420
QY 2744 CATTGCGCTCTTCTTTTCAATTTTCCGACCTGATCTTCCATGAGATGAGGTGGTGT 2803
Db 421 CATTGCGCTCTTCTTTTCAATTTTCCGACCTGATCTTCCATGAGATGAGGTGGTGT 480
QY 2804 TGGCATTGACGAGTGGTGGAGGAATGAACAGTTCTGGGTCAATGGAGGTATCTCTGCACA 2863
Db 481 TGGCATTGACGAGTGGTGGAGGAATGAACAGTTCTGGGTCAATGGAGGTATCTCTGCCCA 540
QY 2864 TCTGTTTGGCGCTCTTTCAGGGTCTTCTGAAGGTGCTTGGCCGATTCGACACCAACTTCAC 2923
Db 541 TCTGTTTGGCGCTCTTTCAGGGTCTTCTGAAGGTGCTTTCGAGGTATCGAGTATCGACACCAACTTCAC 600
QY 2924 TGTCACTCAAAGGCTAATGACGAAGAGGCGACTTTGCTGAGCTTACATGTTCAAGTG 2983
Db 601 TGTCACTCAAAGGCTAATGATGAAGGCGACTTTGCTGAGCTTACATGTTTCAAGTG 660
QY 2984 GACGACGCTTCTCATCCCTCCGACGACCAATTTTGATCATTA 3024
Db 661 GACGACGCTTCTTATCCCTCCGACGACCAATTTTGATCATTA 701
```

Search completed: August 23, 2004, 07:56:30
Job time : 8691 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2004, 01:08:04 ; Search time 25 Seconds
(without alignments)
2230.242 Million cell updates/sec

Title: US-09-900-237A-30
Perfect score: 5778
Sequence: 1 MDGDADALKSGRHGADVQC.....VDPFTTLAGPNIQTGGINC 1080

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2.6/prodata/2/iaa/5A_COMB.pep:*
2: /cgn2.6/prodata/2/iaa/5B_COMB.pep:*
3: /cgn2.6/prodata/2/iaa/6A_COMB.pep:*
4: /cgn2.6/prodata/2/iaa/6B_COMB.pep:*
5: /cgn2.6/prodata/2/iaa/PCUS_COMB.pep:*
6: /cgn2.6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4499	77.9	1065	4	US-09-221-013A-10
2	4455.5	77.1	881	3	US-08-960-048-8
3	4455.5	77.1	881	4	US-09-838-586-8
4	3963.5	68.6	1081	4	US-09-221-013A-6
5	3959.5	68.5	1081	4	US-09-221-013A-12
6	3633.5	62.9	1084	4	US-09-221-013A-8
7	3477	60.2	974	3	US-08-960-048-6
8	3477	60.2	974	4	US-09-838-586-6
9	2811	48.7	685	3	US-08-960-048-7
10	2811	48.7	685	4	US-09-838-586-7
11	2629.5	45.5	629	4	US-09-221-013A-2
12	1756.5	30.4	547	4	US-09-221-013A-14
13	282	4.9	57	4	US-08-812-008-48
14	268	4.6	693	3	US-08-960-048-11
15	268	4.6	693	4	US-09-838-586-11
16	255.5	4.4	861	3	US-08-960-048-12
17	255.5	4.4	861	4	US-09-838-586-12
18	251.5	4.4	946	4	US-09-489-039A-11056
19	238.5	4.1	3031	1	US-07-689-008-2
20	237.5	4.1	756	3	US-08-960-048-10
21	237.5	4.1	756	4	US-09-147-236-3
22	237.5	4.1	756	4	US-09-522-474-3
23	237.5	4.1	756	4	US-09-838-586-10
24	236.5	4.1	738	4	US-09-543-681A-7528
25	230	4.0	883	4	US-09-543-681A-6947
26	210.5	3.6	703	4	US-09-489-039A-10820
27	207	3.6	723	3	US-08-960-048-9

28	207	3.6	723	4	US-08-541-939-2	Sequence 2, Appli
29	207	3.6	723	4	US-09-838-586-9	Sequence 9, Appli
30	207	3.6	723	5	PCT-US91-01726-4	Sequence 4, Appli
31	205	3.5	45	4	US-08-812-008-47	Sequence 47, Appli
32	187.5	3.2	620	4	US-09-489-039A-7689	Sequence 7689, Ap
33	133.5	2.3	904	4	US-09-252-991A-19257	Sequence 19257, A
34	115.5	2.0	501	4	US-09-328-352-5227	Sequence 5227, Ap
35	114	2.0	357	1	US-08-119-773-4	Sequence 4, Appli
36	112	1.9	1957	4	US-08-669-656A-8	Sequence 8, Appli
37	111	1.9	1957	4	US-08-669-656A-2	Sequence 2, Appli
38	110	1.9	797	4	US-09-417-197-141	Sequence 141, App
39	108.5	1.9	419	4	US-09-328-352-7089	Sequence 7089, Ap
40	108	1.9	1167	1	US-08-485-568A-6	Sequence 6, Appli
41	108	1.9	1167	2	US-08-590-554A-6	Sequence 6, Appli
42	108	1.9	1167	2	US-09-184-223-6	Sequence 6, Appli
43	108	1.9	1168	1	US-08-620-717A-9	Sequence 9, Appli
44	107.5	1.9	2132	4	US-08-669-656A-6	Sequence 6, Appli
45	106	1.8	357	1	US-08-119-773-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-221-013A-10
; Sequence 10, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Feng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1065
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-10

Query Match	Best Local Similarity	Score	DB	Length	Indels	Gaps
Matches 835; Conservative	77.9%;	Score 4499;	DB 4;	Length 1065;		
19	COICADGLGTTLDGVTACDVPVPCVYERKEGTQACLOCTKYKRGSPAIR	78				
20	COICSDNVGKVGDRFVACDLCSPVPCVYERKDNQSCPOCKTRYKRLGSPAIP	79				
79	GEEGDDTDADPGS-DFNYPASGTEDQKQIADRMKSRMNTGGSGNVGHPKYDSGEIGLS	137				
80	GDKDEDGLADSGTVEFNYP-----QKEKISERMLGHWLTKGGEEMEPQYDK-----	127				
138	KYDSGEIPRGVPSVTNSQ-MSGEIPGASPDHMMSPGTNRRAPPFY-VNHSFNPSPRE	195				
128	-----EVSHNHLPLRTSRQDTSGEFSAASPERLSVSSTIAGGKRLPYSSDVNQSPNRRIV	182				
196	FSGSIGNVANKERYDGMKMDKGAIPMTNGTSTAPSEGRATDIDASTETNMDALLND	255				
183	DPVGLGNVANKERYDGMKMDKGAIPMTNGTSTAPSEGRATDIDASTETNMDALLND	238				
256	ETROPLSRKVPFIASSKINPYRMVILRLVLSIFLHYRLTNPNVFNAYPLWLLSVICEIWF	315				
239	EARQLLSKVSIPSSRINPYRMVILRLVLSIFLHYRLTNPNVFNAYPLWLLSVICEIWF	298				
316	ALSWILDQFPKFFINRETYLDRLALRYDREGPSQLAAVDIFVSTVDPLKEPPIVTANT	375				

Db 299 ALSWLDDQPKPPVNNRETYLDRLALRYDREGEPSQLAANDVFSVTDPLKPEPLVTANT 358
QY 376 VLSILAVDYPVKVSCYVSDDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPFPYC 435
Db 359 VLSILAVDYPVKVSCYVSDDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPFPYC 418
QY 436 QKIDYLDKQVQSFVKDRAMKREYEEFKIRINALVSKALKVPEEGWIMODCTPWPENNT 495
Db 419 AKIDYLDKQVQSFVKDRAMKREYEEFKIRINALVSKALKVPEEGWIMODCTPWPENNT 478
QY 496 ROHPGMIQVFLGSGGLDTEGNEPRLVYVSRREKRPFGQHHKAGAMNALVRYSVAVLTNG 555
Db 479 GDEPGMIQVFLGSGGLDTEGNEPRLVYVSRREKRPFGQHHKAGAMNALVRYSVAVLTNG 538
QY 556 QYMLNDCDHYINNSKAVREAMCFMDPNLGPQVVCYVFPQRPDGDIDRNDRYANRNVTF 615
Db 539 PFILNDCDHYINNSKALREAMCFMDPNLGPQVVCYVFPQRPDGDIDRNDRYANRNVTF 598
QY 616 DINLRLDGLDQGVYVGTGVFNRTALGYEPPKA--KKPGFLASL--CGGKKKASKSKR 673
Db 599 DINLRLDGLDQGVYVGTGVFNRTALGYEPPKA--KKPGFLASL--CGGKKKASKSKR 658
QY 674 SSDKXSNKHVDSSVPVFNLEIEEGVEGAGFDDKSLMSQMSLEKRFQGSAAVASTL 733
Db 659 ESKKXSGRHTDSTVPVFNLEIEEGVEGAGFDDKSLMSQMSLEKRFQGSAAVASTL 718
QY 734 MEYGGVPQSSTPESLKEAHLVHISCGYEDKSEWTEIGWYGSVTEIDLTGFKMHARGWR 793
Db 719 MENGVPSPATPENFLKEAHLVHISCGYEDKSDMGMEIGWYGSVTEIDLTGFKMHARGWR 778
QY 794 SVTCMPKPAFGKSAPINLSDRLNQVLRWALGVSVEILFSRHCPWYGYGRLKFLERFAY 853
Db 779 SIYCMPLKPAFGKSAPINLSDRLNQVLRWALGVSVEILFSRHCPWYGYGRLKFLERFAY 838
QY 854 INTTYPLTSLPILVYCLPAICLLTGKFMPEISNLSIAFWIALFSLFISFATGILEMRWS 913
Db 839 VNTTYPLTSLPILVYCLPAICLLTGKFMPEISNLSIAFWIALFSLFISFATGILEMRWS 898
QY 914 GVGIDWNRNEQFVIGGISAHLFAVFGQLKVLVLAGIDTNTFTVTSKANDDEGDFAEIYLF 973
Db 899 GVGIDWNRNEQFVIGGISAHLFAVFGQLKVLVLAGIDTNTFTVTSKANDDEGDFAEIYLF 958
QY 974 KWTLLPPTILLINMGVVGAGTSYAINSGYQSWGPLFGKLPFAFWIVHLYPFLKGLM 1033
Db 959 KWTLLPPTILLINMGVVGAGTSYAINSGYQSWGPLFGKLPFAFWIVHLYPFLKGLM 1018
QY 1034 GRQNRTPITVWVAVLLASIFSLWVRVDPPTTRLAGPNIQTCCGNC 1080
Db 1019 GRQNRTPITVWVAVLLASIFSLWVRVDPPTTRLAGPNIQTCCGNC 1065

RESULT 2

US-08-960-048-8
; Sequence 8, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; CURRENT FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Oryzae sativa
US-08-960-048-8

Query Match 77.1%; Score 4455.5; DB 3; Length 881;
Best Local Similarity 93.4%; Pred. NO. 0;
Matches 823; Conservative 39; Mismatches 18; Indels 1; Gaps 1;
QY 201 GNVANKERVYDGMKQDKGAIPTWNGTSTAPSEGRAATDIDASTENMEDALINDETROP 260
Db 1 GNVANKERVYDGMKQDKGAIPTWNGTSTAPSEGRAATDIDASTENMEDALINDETROP 60
QY 261 LSRKVPFIASSKINPYRMWIVLRLVLSIFLHVLRLNTPVRNAYPLWLSVICETWFAISWI 320
Db 61 LSRKVPFIASSKINPYRMWIVLRLVLSIFLHVLRLNTPVRNAYPLWLSVICETWFAISWI 120
QY 321 LQOPKWFPIINRETYLDRLALRYDREGEPSQLAANDVFSVTDPLKPEPLVTANTVLSIL 380
Db 121 LQOPKWFPIINRETYLDRLALRYDREGEPSQLAANDVFSVTDPLKPEPLVTANTVLSIL 180
QY 381 AVDYPVKVSCYVSDDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPFPYCQKIDY 440
Db 181 AVDYPVKVSCYVSDDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPFPYCQKIDY 240
QY 441 LKDKVQPSFVKDRAMKREYEEFKIRINALVSKALKVPEEGWIMODCTPWPENNRDHPG 500
Db 241 LKDKVQPSFVKDRAMKREYEEFKIRINALVSKALKVPEEGWIMODCTPWPENNRDHPG 300
QY 501 MIQVFLGSGGLDTEGNEPRLVYVSRREKRPFGQHHKAGAMNALVRYSVAVLTNGQYMLN 560
Db 301 MIQVFLGSGGLDTEGNEPRLVYVSRREKRPFGQHHKAGAMNALVRYSVAVLTNGQYMLN 360
QY 561 LDCDHYINNSKAVREAMCFMDPNLGPQVVCYVFPQRPDGDIDRNDRYANRNVTFDINLR 620
Db 361 LDCDHYINNSKAVREAMCFMDPNLGPQVVCYVFPQRPDGDIDRNDRYANRNVTFDINLR 420
QY 621 GLDGIQGVYVGTGVFNRTALGYEPPKA--KKPGFLASL--CGGKKKASKSKSDDKKK 679
Db 421 GLDGIQGVYVGTGVFNRTALGYEPPKA--KKPGFLASL--CGGKKKASKSKSDDKKK 480
QY 680 SNKHVDSSVPVFNLEIEEGVEGAGFDDKSLMSQMSLEKRFQGSAAVASTLMEYGGV 739
Db 481 SNKHVDSSVPVFNLEIEEGVEGAGFDDKSLMSQMSLEKRFQGSAAVASTLMEYGGV 540
QY 740 PQSTPESLKEAHLVHISCGYEDKSEWTEIGWYGSVTEIDLTGFKMHARGWRSVYCM 799
Db 541 PQSTPESLKEAHLVHISCGYEDKSEWTEIGWYGSVTEIDLTGFKMHARGWRSVYCM 600
QY 800 KRPAPKGSAPINLSDRLNQVLRWALGVSVEILFSRHCPWYGYGRLKFLERFAYINTIY 859
Db 601 KRPAPKGSAPINLSDRLNQVLRWALGVSVEILFSRHCPWYGYGRLKFLERFAYINTIY 660
QY 860 PLTSLPILVYCLPAICLLTGKFMPEISNLSIAFWIALFSLFISFATGILEMRWSGVGDE 919
Db 661 PLTSLPILVYCLPAICLLTGKFMPEISNLSIAFWIALFSLFISFATGILEMRWSGVGDE 720
QY 920 WWRNEQFVWIGGISAHLFAVFGQLKVLVLAGIDTNTFTVTSKANDDEGDFAEIYMFKWTLL 979
Db 721 WWRNEQFVWIGGISAHLFAVFGQLKVLVLAGIDTNTFTVTSKANDDEGDFAEIYMFKWTLL 780
QY 980 IPTTILLINMGVVGAGTSYAINSGYQSWGPLFGKLPFAFWIVHLYPFLKGLMGRQNT 1039
Db 781 IPTTILLINMGVVGAGTSYAINSGYQSWGPLFGKLPFAFWIVHLYPFLKGLMGRQNT 840
QY 1040 PTIVVWVAVLLASIFSLWVRVDPPTTRLAGPNIQTCCGNC 1080
Db 841 PTIVVWVAVLLASIFSLWVRVDPPTTRLAGPNIQTCCGNC 881

RESULT 3

US-09-838-586-8
; Sequence 8, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE OF INVENTION: Sequences

```
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Oryzae sativa
US-09-838-586-8

Query Match      77.1%; Score 4455.5; DB 4; Length 881;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 823; Conservative 39; Mismatches 18; Indels 1; Gaps 1;

QY 201 GNVAKERVVDGKMKQDKGAIPMTNGTSTAPSEGRAATDIDASTEYNMEDALLNDETROP 260
Db 1 GNVAKERVVDGKMKQDKGAIPMTNGTSTAPSEGRGVGDIDASTDYNNEDALLNDETROP 60
QY 261 LSRKVPASCKNPYRMVILRLVLSIFLHYRLTNPNVFNAYPLMLLSVICIWFALSWI 320
Db 61 LSRKVPFSSRNPNYRNVLRLVLSIFLHYRLTNPNVFNAYPLMLLSVICIWFALSWL 120
QY 321 LDQFPKWPPIRETLYDLRALAYDREGPSOLAANDIFVSTVDPLKEPPIVANTVLSIL 380
Db 121 LDQFPKWPPIRETLYDLRALAYDREGPSOLAANDIFVSTVDPMKEPPIVANTVLSIL 180
QY 381 AVDPVVDKVCYSVDDGASMLTFDALAETSEFARKWVPVKYDIEPRAPEYFVQCKIDY 440
Db 181 AVDPVVDKVCYSVDDGASMLTFDALAETSEFARKWVPVKYDIEPRAPEYFVQCKIDY 240
QY 441 LKDKVQPSFVKDRRAMKREYEFKIRINALVSKALKVPEEGIMQDGPWPNGNTRDHPG 500
Db 241 LKDKVHPSFVKDRRAMKREYEFKIRINGLVAQKVPEEGIMQDGPWPNGNTRDHPG 300
QY 501 MIQVFLHSGGLDTSGNELPLRVLYVSREKPGFQHHKAGANALVRVSAVLNQCYYMLN 560
Db 301 MIQVFLHSGGLDTSGNELPLRVLYVSREKPGFQHHKAGANALVRVSAVLNQCYYMLN 360
QY 561 LDCDHYINNSKAVREAMCFMDPNLGPVCYVQFPQRFQDGDINDRYANRNTVFEDINLR 620
Db 361 LDCDHYINNSKALREAMCFMDPNLGRSVYVQFPQRFQDGDINDRYANRNTVFEDINLR 420
QY 621 GLDGIQGPVYVGTGCVFNRTALYGEYPPIKAKKPG-FLASLCGKKKASKSKSSDKKK 679
Db 421 GLDGIQGPVYVGTGCVFNRTALYGEYPPIKQKKGSFLSSLCGRKKASKSKSSDKKK 480
QY 680 SNKHVDSVPVFNLEDEIEGVEGAGFDEKSVLMSQMSLEKRFQGSAAFPVASTLMEYGV 739
Db 481 SNKHVDSVPVFNLEDEIEGVEGAGFDEKSVLMSQMSLEKRFQGSAAFPVASTLMEYGV 540
QY 740 PQSSTPESLLKEAHIVISCGYEDKSEWTEIGWIYGSVTEIDITGFKMHARGWSVYCMP 799
Db 541 PQSAPTESLLKEAHIVISCGYEDKTEWTEIGWIYGSVTEIDITGFKMHARGWSVYCMP 600
QY 800 KRPAFKGAPINLSDRLNQVLRLWALGSVEILLFSRHCPWYGYGRLKFLERFAYINTIY 859
Db 601 KRPAFKGAPINLSDRLNQVLRLWALGSVEILLFSRHCPWYGYGRLKFLERFAYINTIY 660
QY 860 PLTSLPLAVYCTLPAICLLTGKFMPEJSNLASIMFIALFSIFATGILEMRWSVGIDE 919
Db 661 PLTSLPLAVYCTLPAICLLTGKFMPEJSNLASIMFIALFSIFATGILEMRWSVGIDE 720
QY 920 WWRNEQFVIGISAHLPFAVQGLKVLAGIDTNFTVTSKANDERGDFAELVMPKWTLL 979
Db 721 WWRNEQFVIGISAHLPFAVQGLKVLAGIDTNFTVTSKANDERGDFAELVMPKWTLL 780
QY 980 IPPTTILINNVGVVAGTSYAINSQCWGLFGKLPFAFWVIVHLYPFLKGLMGRQNR 1039
Db 980 IPPTTILINNVGVVAGTSYAINSQCWGLFGKLPFAFWVIVHLYPFLKGLMGRQNR 1039

781 IPPTTILINNVGVVAGTSYAINSQCWGLFGKLPFAFWVIVHLYPFLKGLMGRQNR 840
1040 PTIVIVWVLLASIFSLWVRDPFTRLAGPNQTCGNC 1080
841 PTIVVWVAILLASIFSLWVRDPFTRTRVGTQTCGNC 881

RESULT 4
US-09-221-013A-6
; Sequence 6, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-6

Query Match      68.6%; Score 3963.5; DB 4; Length 1081;
Best Local Similarity 68.0%; Pred. No. 0;
Matches 733; Conservative 138; Mismatches 168; Indels 39; Gaps 15;

QY 2 DGDALAKSGRHGAGDVCGIICADGLGTTLDGVDFTACDVCRPCPCVEHREKGTQAC 61
Db 25 DGGTKPLKNMN--CGICQICGDDVGLAETGVDVFACNECAPVCRPCVEYERKDGTOCC 81
QY 62 LQCKTKYKRRGSPAIRGEGDDTDADDGSDPNYPASGTEDQOKIADRMRSWRNMTGS 121
Db 82 PQCKTRFRHRGSPVEGEDDEDVDDIENEFYAQKANKARHQHGEFSS----- 133
QY 122 GNVGHPKYDSIEIGLSKYD---SGEIPRGYVPSVNTSQMSGEIPGASPDHMMWSPGTNIS 178
Db 134 ----SSRHSQPIPLLTHTGHTVSGEIRTPDQSVRTT--SGPL-GPSDRNATSSP--YID 184
QY 179 FRAPPYVNHSPNPSREF-SGSIGNVAMKERVDMKODKGAIPMTNGTSTAPSEGRA 237
Db 185 PRQVPV--RIVDPKOLNSYGLGNVDWKERVGKWKQKKNMLQMTG----KYHEGRGG 238
QY 238 TDIDASTEYNMEDALLNDETROPILSRKVPPIASSKINPYRMVILRLVLSIFLHYRLTNP 297
Db 239 -EIE-GTSGNGEELQMADDTRLPMRSRVVPISSRLTPYRVVILRLILCFLLQVTRTHP 296
QY 298 VRNAYPLMLLSVICIWFALSWILDQPPKWPPIRETLYDLRALYDREGPSQLAAVDI 357
Db 297 VKNAYPLMLTSVICIWFALSWILDQPPKWPPIRETLYDLRALYDREGPSQLVPVDV 356
QY 358 FVSTVDPLKEPPIVANTVLSILAVDPVVKVSCVSDGASMLTTFDALAETSEFARKW 417
Db 357 FVSTVDPLKEPPIVANTVLSILSDVPVVDKVCYVSDGASMLTTFSLSETAFKWKV 416
QY 418 PFVKYDIEPRAPEYFCQKIDYLDKQVPSFVKDRRAMKREYEEFKIRINALVSKALKV 477
Db 417 PFCKENIEPRAPEYFAQKIDYLDKQLQPSFVKERRAMKREYEEFKIRINALVAKQKI 476
QY 478 PEEGWIMQDGPWPNGNTRDHPGMIQVFLHSGGLDTGEGNELPLRVYVSREKRFQHHK 537
Db 477 PEEGWTMQDGPWPNGNTRDHPGMIQVFLHSGGLDTGEGNELPLRVYVSREKRFQHHK 536
QY 538 KAGAMNALVRVSAVLNQCYYMLNLDCCHYINNSKAVREAMCFMDPNLGPVCYVQFPQ 597
```

Db 537 KAGAMNLRVSAVLNGLNGAYLLNVDCHYFNNSKALKKAMCFMMDPAIGKKCYVFPQR 596
QY 598 FGDIDRNDYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYPPPIKAK--KPG 655
Db 597 FGDIDLHDYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYPPPIKAK--KPG 656
QY 656 FLA-SLCGGKKKASKKSSDKKKXKHNKVDSSVPFNLEDBEGVEGAGFDDEKSVLMS 714
Db 657 IIVKSCGSRKKGKSKYKRRGINRSDSNAPLFNMEDIDEGFE--GYDDERSILMS 714
QY 715 QMSLEKRFQGSAAFAVASTLMYEGVQOSSTPESILKEAHHVISCYEDKSEWGTGKIWIY 774
Db 715 QRSVEKRFQGSAAFAVASTLMYEGVQOSSTPESILKEAHHVISCYEDKSEWGTGKIWIY 774
QY 775 GSVTEDLTGFKMHARGWRSVYCMKRPAPKFGSAPINLSRDLNQLRWALGSEVILFSRH 834
Db 775 GSVTEDLTGFKMHARGWRSVYCMKRPAPKFGSAPINLSRDLNQLRWALGSEVILFSRH 834
QY 835 CPLWYGYGRLKFLERFAYINTIYPLTSLPLAVYCLIPALICLLTGKFMPEISNLASIW 894
Db 835 CPIWYGYGRLKFLERFAYINTIYPLTSLPLAVYCLIPALICLLTGKFMPEISNLASIW 894
QY 895 FIALFISIFATGILEMRWSGVIGDEWNRNQFWVIGGISAHLPFAVFGGLKVLGIDTNF 954
Db 895 FILLFISIAVTGILEMRWSGVIGDEWNRNQFWVIGGISAHLPFAVFGGLKVLGIDTNF 954
QY 955 TVTSKANDEGDFAEIYFKWMTLLIPPTTILINMVGVVAGTSYAINSGYQSWGLFGK 1014
Db 955 TVTSKANDEGDFAEIYFKWMTLLIPPTTILINMVGVVAGTSYAINSGYQSWGLFGK 1014
QY 1015 LFFAFWVIAHLVPLKGLGRQNTPTIVIVWVLLASIFSLWVRVDPFTRLAGPN 1072
Db 1015 LFFALWVIAHLVPLKGLGRQNTPTIVIVWVLLASIFSLWVRVDPFTRLAGPN 1070

RESULT 5

US-09-221-013A-12
; Sequence 12, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221.013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-12

Query Match 68.5%; Score 3959.5; DB 4; Length 1081;
Best Local Similarity 67.9%; Pred. No. 0;
Matches 732; Conservative 138; Mismatches 169; Indels 39; Gaps 15;

QY 2 DGDADALKSGRHGADVCQICADGLGTLTLDGVFTACDVCRPPVCRPCYEHKEGTQAC 61
Db 25 DGGTKPLKNWN--GQICQICGDDVGLAETGDFVACNECAFFVCRPCYEHKEGTQAC 81
QY 62 LQCKTKYKRRGSAIRGEEDDDTDADGSDFNYPASTEDQOKIADRMESWRMTGGS 121
Db 82 PQCKTRFRHRGSRVDEGDEDDVDVDIENEFNVAQGANRKHQRHGEFSS----- 133

QY 122 GNVGHPKYDSBIGLSKYD---SGEIPRGYVSVTNQSOMSGEIPGASPDHMMSPGNTIS 178
Db 134 ---SRHESQPIPLLTHGTVSGEIRTPDTQSVRTT--SGPL-GPSDRNAISSP--YID 184
QY 179 RRAPPFYVNHSPNSREP--SGSIGNAVWERVDGWMKQDKGAIPMTNGTSTIAPSEGRA 237
Db 185 PRQFPVPV--RIVDPKSDILNSYGLGNVDKERVGKWKQKMLQMTG---KYHGGKGG 238
QY 238 TUIDASTEYNMEDALLNDETROPISRKVPPIASSKINPVMVIVLRVLWLSIFLHYRLTNP 297
Db 239 -EIE-GTGNGBELQWADTRLPMSRVPIPSRLTPYRVVILRLILLCFFLOYRTHP 296
QY 298 VRNAYPLWLLSVEICEIWFALSWILDQFPKWFPIINRETYLDRALRYDREGEPSQLAAVDI 357
Db 297 VKNAYPLWLLSVEICEIWFALSWILDQFPKWFPIINRETYLDRALRYDREGEPSQLAAVDI 356
QY 358 FVSTVDPLEKPPVITANTVLSILAVDYPVDKUSCYVSDDGASMLTDAEAETSEFARKW 417
Db 357 FVSTVDPLEKPPVITANTVLSILAVDYPVDKUSCYVSDDGASMLTDAEAETSEFARKW 416
QY 418 PFVKYDIDIPRAPEFYFCQKIDYLDKVPSPFVKDRRAMKREYEEFKIRINALVSKALKV 477
Db 417 PFCKKNIERAPREFYFAQKIDYLDKVPSPFVKDRRAMKREYEEFKIRINALVSKALKV 476
QY 478 PEEGWMQDGTWPNGNTRDHPQMIQVFLHSGGLDTEGNEPLRLVTVSREKPGFQHHK 537
Db 477 PEEGWMQDGTWPNGNTRDHPQMIQVFLHSGGLDTEGNEPLRLVTVSREKPGFQHHK 536
QY 538 KAGAMNLRVSAVLNGLNGAYLLNVDCHYFNNSKALKKAMCFMMDPAIGKKCYVFPQR 597
Db 537 KAGAMNLRVSAVLNGLNGAYLLNVDCHYFNNSKALKKAMCFMMDPAIGKKCYVFPQR 596
QY 598 FDGIDRNDYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYPPPIKAK--KPG 655
Db 597 FDGIDLHDYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYPPPIKAK--KPG 656
QY 656 FLA-SLCGGKKKASKKSSDKKKXKHNKVDSSVPFNLEDBEGVEGAGFDDEKSVLMS 714
Db 657 IIVKSCGSRKKGKSKYKRRGINRSDSNAPLFNMEDIDEGFE--GYDDERSILMS 714
QY 715 QMSLEKRFQGSAAFAVASTLMYEGVQOSSTPESILKEAHHVISCYEDKSEWGTGKIWIY 774
Db 715 QRSVEKRFQGSAAFAVASTLMYEGVQOSSTPESILKEAHHVISCYEDKSEWGTGKIWIY 774
QY 775 GSVTEDLTGFKMHARGWRSVYCMKRPAPKFGSAPINLSRDLNQLRWALGSEVILFSRH 834
Db 775 GSVTEDLTGFKMHARGWRSVYCMKRPAPKFGSAPINLSRDLNQLRWALGSEVILFSRH 834
QY 835 CPLWYGYGRLKFLERFAYINTIYPLTSLPLAVYCLIPALICLLTGKFMPEISNLASIW 894
Db 835 CPIWYGYGRLKFLERFAYINTIYPLTSLPLAVYCLIPALICLLTGKFMPEISNLASIW 894
QY 895 FIALFISIFATGILEMRWSGVIGDEWNRNQFWVIGGISAHLPFAVFGGLKVLGIDTNF 954
Db 895 FILLFISIAVTGILEMRWSGVIGDEWNRNQFWVIGGISAHLPFAVFGGLKVLGIDTNF 954
QY 955 TVTSKANDEGDFAEIYFKWMTLLIPPTTILINMVGVVAGTSYAINSGYQSWGLFGK 1014
Db 955 TVTSKANDEGDFAEIYFKWMTLLIPPTTILINMVGVVAGTSYAINSGYQSWGLFGK 1014
QY 1015 LFFAFWVIAHLVPLKGLGRQNTPTIVIVWVLLASIFSLWVRVDPFTRLAGPN 1072
Db 1015 LFFALWVIAHLVPLKGLGRQNTPTIVIVWVLLASIFSLWVRVDPFTRLAGPN 1070

RESULT 6

US-09-221-013A-8
; Sequence 8, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.

```

; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucon
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1084
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-8

Query Match          62.9%; Score 1633.5; DB 4; Length 1084;
Best Local Similarity 61.7%; Pred. No. 0;
Matches 679; Conservative 171; Mismatches 192; Indels 59; Gaps 16;

QY 1 MDGDADA-LKSGRHGAGDVQICADGLGTTLDGDVFTACDVCRFPVPCPYEHERKEGTQ 59
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
20 INADESARISQVLSGQTQICGDEITELTVSSSELFVACNECAFPVCRPCPYEYERREGNQ 79
QY 60 ACLOCKTKYKHHGSGPAIRGEGDDTDADGSDNPNYPASGTEDQOKIADMRSWRNMTG 119
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
80 ACPOCKTKYKRIKGSPRV---DGDDEEBEEDIDDLLEYDFDHGMD--PEHAAEAALSSRLNTG 135
QY 120 GSGNVGHFKYDSDGIGLSKYDSGIBPGYVPSVTNSQMSGEIPGASDPDHMMGP--TGN 177
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
136 RGGLDSP-----PQSQIPLLTYCDEDADM--YSRHALIYVPSTGYG 176
QY 178 SRRAPFPYVNHSPNS-----REFSGSIGNVAWKERVQGWKMQ--DKGAIPMTNG 226
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 NRVPYPAFTSSAPPQARSVPVKDIAEY--GYGSVAWKDRMEVWKRQGEKLVQIKHEG 234
QY 227 TSIAPSGRAATDIDASTEYNMEDALNDSTROPLSKVPDIASKINPYRMVILRLVL 286
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
235 ----GNNGRSNDDD---ELDDPMPMDSGROPLSKLPIRSRINPYRMLLCLRLAIL 287
QY 287 SIFLHYRLTFVRNAYPLWLLSICEIWFALSWILDQPKWFFPINRETYLDRLLALRYDRE 346
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
288 GLFFHYKILLHPVNDAYGLMLTSLVICEIWFVAVSWILDQPKWYPIERETYLDRLSLAYEKE 347
QY 347 GEFSQLAAVDIFVSTVDPLKEPIVIVTANTVLSILAVDVPVDKVCYVSSDDGASMLTFDAL 406
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
348 GKESGLAPVDVYSTVDPLKEPLITANTVLSILAVDVPVDKVCYVSNNGAAMLTFEAL 407
QY 407 AETSEFARKWVPVKYDIEPRAPERYFCOKIDYKDKVQPSFVKORRAKMRKRYEFKIR 466
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
408 SDTADFAIKWVPCKKFNIEPRAPETFSQMDYLKKNVHPAFVREKRRAKMRKRYEEFKVK 467
QY 467 INALVSKALKVPEEGIMQDGTWPNGNTRDHPGMIQVFLGSHGGGLDTEGNELPRLYVVS 526
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
468 INALVATAQVPEERWTMQDGTWPNGNVNVDHPGMIQVFLGHSVSRDTCGNELPRLYVVS 527
QY 527 REKRPGFQHHKKAGAMNALVRVSAVLNNGQVYMLNDCDHVYNNNSKAVREAMCFMDPNLG 586
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
528 REKRPGFDHKKKAGAMNSLIRVSAVLNSAPYLLNVDCHDYINNSKATIRESCMCFMDPQSG 587
QY 587 PQVCYVQFQRFQDGI DRNDRYANRNTVFFDINLRGLDGIQPVYVGTGCVFNRTALVGYE 646
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
588 KKVCYVQFQRFQDGI DRHDYSNRNVVFIDNKKGLDGIQSPIYVGTGCVFRKQALYGF 647
QY 647 PPTIAKKKPGFLAS-----LCGGKKKVAASKKRSDDKKSKNKHVDSVVPVFNLEDEEG 699
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
648 APKKKKKPPGKTCNCWPKWCCLCCGLNRK--KSKTKATDKKNTK--ETSKQIHALENVDEG 703
QY 700 VEGAGFDDEKSVLMSQMSLEKRGQSNAAFVASTILMEYGGVQPSSTESLILKEAIIHVISG 759
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
704 VIVPVSNVKRSBATQLKLEKFGQSPVFVASAVLQNGGVPRNAPSACILREAIQVISC 763

```

```
QY 377 LSTLAVDPVVKVSCVSDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPFVFCQ 436
Db 274 LSTLADVPVVKVSCVSDGASMLTFESLVTADFAKWKVPFCKFESIEPRAPFVFSQ 333
QY 437 KIDYLDKQVPSFVKORRAMKREYEFKIRINALVSKALKVPBEGTMDGTPWPGNNT 496
Db 334 KIDYLDKQVPSFVKERRAMKRDYEVKIRINALVAKAQTPDEGTMQDGTSPWGNPNR 393
QY 497 DHPGMIQVFLHSGGLDTEGNELPRLVYVSREKPGFOHKKAGAMNALVRVSAVLN 556
Db 394 DHPGMIQVFLGSGARDIEGNELPRLVYVSREKPGYQHHKAGAMNALVRVSAVLN 453
QY 557 YMLNLCDDHYNNKSKAVREACFMDPNLGPQVYVQFQORFDGIDRNDRYANRNT 616
Db 454 FIUNLDCDDHYNNKSKAVREACFMDPQVGRDVCYVQFQORFDGIDRSDRYANR 513
QY 617 INLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPFIAKKPGFLASLGGKKKASKRS 676
Db 514 VNMKGLDGIQGPVYVGTGCVFNRAQALYGYGPPSMPSPKSSSSSSSCCCPGKE 573
QY 677 KKSNNKHVDSSVPFNLDEIEEGVEGAGFDD-EKSVLMSQMSLEKRFQSAAFV 735
Db 574 LYRDAKREELDAAIFNLREID-----NYDEYERSMLISQTSFEKTFGLSS 627
QY 736 YGVVQSSPTPESELKEAHHVISCYEDKSEWGTGTEIGWYGSVTEIDILTGF 795
Db 628 NGGVAESANPSTLKEAHHVISCYEBKTAWGKEIGWYGSVTEIDILTGF 687
QY 796 YCMKPRPAFGKSAPINLSRDLNOVLRWALGSVEILFSRHCPPLWYGY-GGR 854
Db 688 YCMPLRPAFGKSAPINLSRDLHQLRWALGSVEILFSRHCPPLWYGYGFGGR 747
QY 855 NTIYPLTSLPLAVCYCLPAICLLTGKFIPEISNLASIWFIALFLSIFATG 914
Db 748 NTIYPTSLPLAYCSLPAICLLTGKFIPTLSNLASVILFLGFLSIIVTAV 807
QY 915 VGIDWWRNEQFWVIGGISAHFLFAVFOGLLKVLAGIDTNTVTSKANDEG 974
Db 808 VSIEDLWRNEQFWVIGGVSAILFAVFOGLKMLAGIDTNTVTAKAAD-DAD 866
QY 975 WTTLLIPPTTILINMVGWVAGTSYAINSGYQSGWPLFGKLPFAFWIVH 1034
Db 867 WTTLLIPPTTILINMVGWVAGSDALNKGYEAWGPLFGKVFVFWILH 926
QY 1035 RQNRTPITIVWAVILLASIFSLWVRVDPPTTRLAGPNI-QTC-GINC 1080
Db 927 RQNRTPITIVWSVLLASVFLWVRINPFPVSTADSTTVSQSLCISIDC 974
```

RESULT 8

```
US-09-838-586-6
; Sequence 6, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-09-838-586-6
```

```
Query Match 60.2%; Score 3477; DB 4; Length 974;
Best Local Similarity 61.5%; Pred. No. 1.5e-318;
Matches 657; Conservative 129; Mismatches 176; Indels 106; Gaps 15;

QY 18 VCOICADGLGTTLDGDFVFTACDVPVCRPCVEHERKEGTQACLOCKTKYKRRGSPAI 77
Db 8 VCHTCGEHVLNVNGEPFFVACHCECNFPICKSCPEYDLKEGRKACLR-----GSPY- 58
QY 78 RGEGBGDDTADGSDPNYPASGTEDQKQIADRMRSWRMNTGSGNVG-HPKYDSIGL 136
Db 59 -----DENLDD-----VEKATGDOSTMAA-----HLNKSQDVIHARHIS--SV 96
QY 137 SKYDSGEIPRGVPSVTNSQMSGEIFGASPDHMMSPGTGNI SRRAPFPYVNHSPNREF 196
Db 97 STLDS-----EM 103
QY 197 SGSIGNVAMKERVDMKMKODKGAIPTWNTGTSIAPSEGRAATDIDASTEYNNMEDALINDE 256
Db 104 AEONGNSIMKNRVSVESWKEKNKKKKPAATT-----KVERAEIIPPEQMEDKAPDA 154
QY 257 TROPLSRKVPPIASSKINPYRMVIVLRLVLSIFELHYRLTNPNVNAVPLMLLSVICI 316
Db 155 S-QPLSTIIPKSRAPYRTVIIIMRLIILGLFFHYRVTPNVPDSAFGLMLTSLVICI 213
QY 317 LSWILDQFPKWFPIINRETVLDRALRYDRGEPSQLAANDIFVSTVDPLKEPPIV 376
Db 214 FSWVLQFPKWFVPIINRETVLDRLSARYEREGETDELAADVFFVSTVDPLKEPPI 273
QY 377 LSTLAVDPVVKVSCVSDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPFVFCQ 436
Db 274 LSTLADVPVVKVSCVSDGASMLTFESLVTADFAKWKVPFCKFESIEPRAPFVFSQ 333
QY 437 KIDYLDKQVPSFVKORRAMKREYEFKIRINALVSKALKVPBEGTMDGTPWPGNNT 496
Db 334 KIDYLDKQVPSFVKERRAMKRDYEVKIRINALVAKAQTPDEGTMQDGTSPWGNPNR 393
QY 497 DHPGMIQVFLHSGGLDTEGNELPRLVYVSREKPGFOHKKAGAMNALVRVSAVLN 556
Db 394 DHPGMIQVFLGSGARDIEGNELPRLVYVSREKPGYQHHKAGAMNALVRVSAVLN 453
QY 557 YMLNLCDDHYNNKSKAVREACFMDPNLGPQVYVQFQORFDGIDRNDRYANRNT 616
Db 454 FIUNLDCDDHYNNKSKAVREACFMDPQVGRDVCYVQFQORFDGIDRSDRYANRNT 513
QY 617 INLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPFIAKKPGFLASLGGKKKASKRS 676
Db 514 VNMKGLDGIQGPVYVGTGCVFNRAQALYGYGPPSMPSPKSSSSSSSCCCPGKE 573
QY 677 KKSNNKHVDSSVPFNLDEIEEGVEGAGFDD-EKSVLMSQMSLEKRFQSAAFV 735
Db 574 LYRDAKREELDAAIFNLREID-----NYDEYERSMLISQTSFEKTFGLSS 627
QY 736 YGVVQSSPTPESELKEAHHVISCYEDKSEWGTGTEIGWYGSVTEIDILTGF 795
Db 628 NGGVAESANPSTLKEAHHVISCYEBKTAWGKEIGWYGSVTEIDILTGF 687
QY 796 YCMKPRPAFGKSAPINLSRDLNOVLRWALGSVEILFSRHCPPLWYGY-GGR 854
Db 688 YCMPLRPAFGKSAPINLSRDLHQLRWALGSVEILFSRHCPPLWYGYGFGGR 747
QY 855 NTIYPLTSLPLAVCYCLPAICLLTGKFIPEISNLASIWFIALFLSIFATG 914
Db 748 NTIYPTSLPLAYCSLPAICLLTGKFIPTLSNLASVILFLGFLSIIVTAV 807
QY 915 VGIDWWRNEQFWVIGGISAHFLFAVFOGLLKVLAGIDTNTVTSKANDEG 974
Db 808 VSIEDLWRNEQFWVIGGVSAILFAVFOGLKMLAGIDTNTVTAKAAD-DAD 866
QY 975 WTTLLIPPTTILINMVGWVAGTSYAINSGYQSGWPLFGKLPFAFWIVH 1034
Db 867 WTTLLIPPTTILINMVGWVAGSDALNKGYEAWGPLFGKVFVFWILH 926
QY 1035 RQNRTPITIVWAVILLASIFSLWVRVDPPTTRLAGPNI-QTC-GINC 1080
```


Db 598 GVVAGSDAINNGYSGWGLFGLKLFAPFWIHLHLPFLKGLMGRQRTTIVVLSILIA 657
QY 1052 SIFSLLWVRVDPFTTRLAGENIQTGCI 1078
Db 658 SIFSLLWVRIDPFLPKQTGPVLKQCGV 684

RESULT 11
US-09-221-013A-2
; Sequence 2, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-2

Query Match 45.5%; Score 2629.5; DB 4; Length 629;
Best Local Similarity 75.2%; Pred. No. 4.6e-239;
Matches 468; Conservative 76; Mismatches 71; Indels 7; Gaps 4;
QY 454 RAMKREYEEFKIRINALVSKALVPERGWINODGTPMGNTRDHPGMIQVFLHSGGLD 513
Db 1 RAMKREYEEFKIRINALVSKALVPERGWINODGTPMGNTRDHPGMIQVFLHSGGLD 60
QY 514 TEGNELPRLVVSREKPGFOHKKKAGAMNALVRVAVLTNGQYMLNLDCHYINNSKAV 573
Db 61 TDGNEPRLVVSREKPGFOHKKKAGAMNARSIRVAVLTNGAYLLNVDCDHYFNNSKAI 120
QY 574 REAMCFMNDNLGPQVYVQPPQPDGIDRNDVANNRVFPDNLRLDGIQGPVYVGT 633
Db 121 KEAMCFMNDNLGPQVYVQPPQPDGIDRNDVANNRVFPDNLRLDGIQGPVYVGT 180
QY 634 GCVENRTAIVGEPIKAK--KPGFLA--SLCGGKKKSKSKSSDDKKSNKHVDSSVPV 690
Db 181 GCCFNRLQALYGYDPLVLTEDLEPNILVKS CGSRKKGSKSKYVEKRGINRSDSNAPL 240
QY 691 PNLIEDIEGVGAGFDDEKSVLMSQMSLEKRFQSGAFAVASTLMEYGVGQVPSSTPESLLK 750
Db 241 PNMEDIDEGFE--GYDDERSILMSQSVKEKFGQSPVFIATFMEQGGIPPTTNPATLLK 298
QY 751 EAIHVISCGYEDKEMGTEIGWYGSVTEIDILTGFMHARGWSVYCMKPRPAKGSAPI 810
Db 299 EAIHVISCGYEDKEMGTEIGWYGSVTEIDILTGFMHARGWSVYCMKPRPAKGSAPI 358
QY 811 NLSRLNQLRWALSGVILSRHCPVWYGGRLKELERFAYINTTIYPLTSLPLLYVC 870
Db 359 NLSRLNQLRWALSGVILSRHCPVWYGGRLKELERFAYINTTIYPLTSLPLLYVC 418
QY 871 ILPAICILITGPIEMPEISNLASIFALFSLFATGILMRWSGVGIDEDWRNRNFWWIG 930
Db 419 ILPAICILITGPIEMPEISNLASIFALFSLFATGILMRWSGVGIDEDWRNRNFWWIG 478
QY 931 GISAHLPVQGLLKVLAGIDNTFTVTSKANDDEGDAELMYFKWTTLLIPPTTILINM 990
Db 479 GTSTHLFVQGLLKVLAGIDNTFTVTSKANDDEGDAELMYFKWTTLLIPPTTILINM 538

QY 991 VGVAGTSYAINSGYQSGWGLFGLKLFAPFWIHLHLPFLKGLMGRQRTTIVVWVLL 1050
Db 539 IGIVAGSYAVNNGYQSGWGLFGLKLFAPFWIHLHLPFLKGLMGRQRTTIVVWVLL 598
QY 1051 ASIFSLLWVRVDPFTTRLAGPN 1072
Db 599 ASIFSLLWVRINPFVD--ANPN 618

RESULT 12
US-09-221-013A-14
; Sequence 14, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-221-013A-14

Query Match 30.4%; Score 1756.5; DB 4; Length 547;
Best Local Similarity 61.5%; Pred. No. 1e-156;
Matches 346; Conservative 66; Mismatches 100; Indels 51; Gaps 13;
QY 2 DGDA--DALKSGRHGAGDVQCICADGLGTTLDGVFTACDVCRFPVPCPYEHEKGTQA 60
Db 23 DGDAAPPKAPGKSVNGVQCICGDTGVGSATGDVFAVCNCAFCVPCPYEHEKGTQA 82
QY 61 CLQCKTKYKHRSGSPAIRGEGDDTDADD--GSDPNP---ASGTEDQKXIADRMRSWM 116
Db 83 CPQCKTRYKHKGCPVQGD--EEDVDLONEHYKHGNGKGP-----WQI 129
QY 117 NTGSGNVGHPKYDSGEIGLSKYDSGEIPRGVPSVTNSQMSGEIPGASPDHMMSPGN 176
Db 130 QRQGE-----DVLSSSRHEQHRIPR---LTSCQQISGEIPDASPDHHSI--- 172
QY 177 ISRRAPPFPYVNHSS-----PNPSREF--SGSIGNVAMKERVDMKMKQDKGAIPTWNGTS 228
Db 173 --RSGTSSYVDPSPVFPVRIVDPSKDLNSYSGINSVDMQERVASWRNKQDKNMQVAN--- 227
QY 229 IAPSEGAATDIDASTYNNMEDALLDETROPKSRKYPKPIASSKINPYRMVTLRLVLSI 288
Db 228 --KYPEARGG--DME--GTGNGEDIQWDDARLPLSRIVIPSNQNLNRYIVILRLILMF 284
QY 289 FLHRLTNPVNRYAPLWLLSVICETWFAISWILQFQPKWFFPINRETYLDRLALRYDREG 348
Db 285 FFQFRVTHPVNRYAPLWLLSVICETWFAISWILQFQPKWFFPINRETYLDRLALRYDREG 344
QY 349 PSQAALVDIFVSTVDPLKEPPIVNTVLSILAVDVPDVKVSCYVSDDGASMLTFDALAE 408
Db 345 PSQAALVDIFVSTVDPLKEPPIVNTVLSILAVDVPDVKVSCYVSDDGASMLTFDALAE 404
QY 409 TSEFARKVVPKVKYDIEPRAPFVFCOKIDYLDKQVPSFVKDRAMKREYEEFKIRIN 468
Db 405 TAEFARKVVPKVKYDIEPRAPFVFCOKIDYLDKQVPSFVKDRAMKREYEEFKIRIN 464
QY 469 ALVSKALVVPBEGTMDQGTWPGNTRDHPGMIQVFLHSGGLDTEGNEPRLVVSRE 528
Db 465 ALVSKALVVPBEGTMDQGTWPGNTRDHPGMIQVFLHSGGLDTEGNEPRLVVSRE 524

QY 529 KRPGFOHHKAGAMNALVRVSAV 551
Db 525 KRPGFOHHKAGAMNALIRVSAV 547

RESULT 13

US-08-812-008-48
; Sequence 48, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustine, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 150.183US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-812-008-48

Query Match 4.9%; Score 282; DB 4; Length 57;
Best Local Similarity 92.9%; Pred. No. 3.8e-19;
Matches 52; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 774 YGSVTEDILTGKMHARGRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEI 829
Db 1 YGSVTEDILTGKMHARGRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEI 56

RESULT 14

US-08-960-048-11
; Sequence 11, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; CURRENT FILING DATE: 1997-10-29

; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Escherichia coli
US-08-960-048-11

Query Match 4.6%; Score 268; DB 3; Length 693;

Best Local Similarity 19.5%; Pred. No. 5.7e-16;
Matches 156; Conservative 83; Mismatches 201; Indels 360; Gaps 29;

QY 263 RKVPIASSKINPYRMVILRLVLSIFLHYRUT-----NPVRNAYPLWLLSVICIEWFA 316
Db 11 RMPGRFSAL-----MLIVLSLTVCRYIWMRYTSTLNWDDPVSIVGLLILFAITYAMIV 66
QY 317 LSWILDQPPKWPINRETYLDRLALRYDEGSPSLAAVDIFVSTVDPLKEPPIVTANTV 376
Db 67 L--VLGYFQVWPLNRQP-----VPLPKDMSLWPS-----VDIFVPTYN---EDLNVVKNTI 113
QY 377 LSLAVDYPVDKVCVSDDGASMLTDFDALETSEFARKWVPFVKYDIEPRAPPEYFQ 436
Db 114 YASLGIDWPCKLNIWLLDDGG----- 135
QY 437 KIDYLDKQVQPSFVKDRAMKREYEEFKIRINALVSKALKVPEEGIMQDGTWFGNNTR 496
Db 136 -----REFRQFAQNVG----- 147
QY 497 DHPGMIQVFLGHSGLDTEGNELPLVYVSRKRGFQHHKAGAMNALVRVSAVLNMQ 556
Db 148 -----VKYIARTT-----HEHAKAGINNALKYA-----KGE 174
QY 557 YMLNLCDHYINNSKAVREAMCFMLDPNLGPOVCYVQPPQRE---DGIDRN---DRYAN 609
Db 175 FVSIFDCDHVPTSRFLQMTMGWFLKE---KQLAMQTHFFHFFSDPPERNLGRFKTIN 230
QY 610 RNTVFDDINLRGLDGIQGVVYVGTGCVFNRTAIYGYEPPIIKAKKFGFLASLGGKKKASK 669
Db 231 EGTLYFGLVQDGNMDWDATFFCGSCAVIRR-----KP----- 262
QY 670 SKKRSSDKKSNKHVDVSPVFNLEDIEGVEGAGFDDKSVLMSQMSLEKRFQGSAAFV 729
Db 263 ----- 262
QY 730 ASTLMEYGVQPSSTPESLLKEAIIHVISCGVEDKSEWGTETIGWIYGSVTEDILTGFKMHA 789
Db 263 ---LDEIGGI-----AVE-----TATEDAHTSLRLHR 286
QY 790 RGWRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCPHLYWYGGSLKFLF 849
Db 287 RGYTSAYN--RIPQAAGLATESLSAHIGQIRIWRARGMVQI--FRLDNPL---TGKGLKFAQ 340
QY 850 RFAYINTIYPLTSLPLAVYCLP-----AICLTGKFIIMP----- 885
Db 341 RLCVYNAMEFELSGIPRLIFLTAPLALLHAYIIYAPALMIALFVLPHMIHSLTNSKI 400
QY 886 -----EISNLASIWFTA---LFLSIFATGILEMRWSGVGIDSWRNEOF--WVI 929
Db 401 QGKVRHSFWSSEIYETVLAWYIAPTLLVALINPHKGFKNVTAKGGGLVE---EEVDWVI 456
QY 930 GGISAHILFAVFQGLLKVLAGIDTNTVTSKANDEBGFPAELMYFKWTLLIPTTIL--I 987
Db 457 S--RPYIFPLVLLNLVGVAVGI-----WRYFYGEPTMLIVV 490
QY 988 INWGV-----VAGTSVAIN 1002
Db 491 VSMVWVFNVLIVLGGAVVS 510

RESULT 15
US-09-838-586-11

```
; Sequence 11, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-838-586-11

Query Match      4.8%; Score 268; DB 4; Length 693;
Best Local Similarity 19.5%; Pred. No. 5.7e-16;
Matches 156; Conservative 83; Mismatches 201; Indels 360; Gaps 29;

QY      263 RKVPIASSKINPYRMVILVLSIFLHVRLT-----NPVRNAVPLMLLSVCEIWFA 316
Db      11 RRMPEGRFSAI-----MLIVLSLTSCRYIWKRYTSTLWDDPVSIVCGLLILLFAITYAMIV 66

QY      317 LSWILQDFPKWPFINRETYLDRLALRYDREGPEPSQLAAVDIFVSTVDPLKEPPIVTANTV 376
Db      67 L--VLGYFQVWELNRQP-----VPLPKDMSLWFS---VDIFVPTYN---EDLNVVKNTI 113

QY      377 LSLIANDYPVDKVCYVSDGASMLTFDALAEISEFARKVPFVKKYDIEPRAPEFYFCQ 436
Db      114 YASLIGIDWPDKDLNIWILDGG-----

QY      437 KIDYLDKQVQPSFKDRRAMKREVEEPIKIRINALVSKALKVPEEGIMQDGTWPNGNTR 496
Db      136 -----REEFRQAQNVG-----

QY      497 DHPGMIQVFLGHSGGLDTEGNELPRLVYVSRKRPQGHKKAGAMNALVRVSAVLNQG 556
Db      148 -----VKYIARTT-----HEHAKAGINNALKYA---KGE 174

QY      557 YMLNLCDDHYNNKSAVREANCFMIDPNLGPQVCYVQFPQRF---DGIDEN---DRYAN 609
Db      175 FVSIFDCDHVPTRSFLQMTWGWFLKE---KQLAMMOTPHHFFSPDPFERNLGRFRKTPN 230

QY      610 RNTVFFDINLRGLDGIQGPVVVGCVFENRTAIYGEPPIKAKKPGFLASLCGGKKKASK 669
Db      231 EGTLYGLVQDGNMDWDATFCGSCAVIRR-----KP-----262

QY      670 SKRSDKKKSNKHVDSSVPVFNLEDIERGVEGAGFDDKSVLMSQMSLEKRFQGSAAFV 729
Db      263 -----

QY      730 ASTIMEYGVYQSTPSSLKEATHVISCYVEDKSEWCTEIGWIYGVSTEDILTFKXHA 789
Db      263 ---LDEITGGI-----AVE-----TVTEAHTSLRLHR 286

QY      790 RGMRSVYCMKPKRPAFKGAPINSLDRNLQVLRWALGSVEILFSRHCPLMWYGGRLKFLF 849
Db      287 RGYTSAYM--RIPQAAGLATESLSAHICQIRWARGWVQI--FRLDNPL---TGKGLAFQ 340

QY      850 RFAYINTTYPLTSLPLVYCILP-----AICLLTGKFIIMP-----885
Db      341 RLCYVNAMFHELSGIPRLIFLTAFLAFLHLLHAYIIYAPALMIALFVLPHMIHASLITNSKI 400

QY      886 -----RISNLASTWFA---LFLSIFATGILEMRWSGVGIDBWNENQF--WVI 929
Db      401 QGKYRHSPWSBIYETVLAWIYAPPTLVALINPHKGNFNTAKGGGLVE-----EYVDWVI 456
```

```
QY      930 GGISAHLFAVFGCLLKVLGIDTNTFTVTSKANDEEGDFAELYMKWTTLLIPPTTL--I 987
Db      457 S--RPYIFLVLLNLVGVAVGI-----WRYFYGPPTMLTVV 490

QY      988 INNVGV-----VAGTSYAIN 1002
Db      491 VSMVMVVFYNLLIVLGGAVAVS 510

Search completed: August 23, 2004, 01:11:16
Job time : 29 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2004, 01:09:20 ; Search time 93 Seconds
(without alignments)
3649.414 Million cell updates/sec

Title: US-09-900-237A-30
Perfect score: 5778
Sequence: 1 MDGDADALKSGRHGAGDVQC.....VDPFTTTLRAGNIQTGGINC 1080

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 314255058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5778	100.0	1080	9	US-09-900-237-30
2	5452.5	94.4	1081	16	US-10-437-963-174885
3	5423.5	93.9	1079	12	Sequence 174885,
4	5423.5	93.9	1079	14	US-10-627-132-22
5	5423.5	93.9	1079	14	US-10-267-459-6
6	5421.5	93.8	1077	12	US-10-209-059-22
7	5421.5	93.8	1077	12	US-10-627-132-10
8	5421.5	93.8	1077	14	US-10-209-059-10
9	5421.5	93.8	1077	14	US-10-160-719-6
10	5261.5	91.1	1043	14	US-10-160-719-30
11	5149	89.1	1076	12	US-10-160-719-50
12	5149	89.1	1076	12	US-10-627-132-14
13	5149	89.1	1076	14	US-10-209-059-14
14	5129	88.8	1119	12	US-10-160-719-58
15	5122.5	88.7	1063	16	US-10-425-114-58605
					Sequence 58605, A
					Sequence 142245,

16	4507	78.0	1065	9	US-09-900-237-33	Sequence 33, Appl
17	4499	77.9	1065	14	US-10-229-193-10	Sequence 10, Appl
18	4455.5	77.1	881	9	US-09-838-539-8	Sequence 8, Appl
19	4031.5	69.8	1085	12	US-10-424-599-249738	Sequence 249738
20	4021.5	69.6	1091	9	US-09-900-237-26	Sequence 26, Appl
21	3993.5	69.1	1074	12	US-10-627-132-46	Sequence 46, Appl
22	3993.5	69.1	1074	14	US-10-209-059-46	Sequence 46, Appl
23	3993.5	69.1	1074	14	US-10-160-719-14	Sequence 14, Appl
24	3993.5	69.1	1074	14	US-10-160-719-22	Sequence 22, Appl
25	3993.5	69.1	1074	14	US-10-160-719-42	Sequence 42, Appl
26	3988	69.0	1075	12	US-10-627-132-2	Sequence 2, Appl
27	3988	69.0	1075	14	US-10-209-059-2	Sequence 2, Appl
28	3988	69.0	1075	14	US-10-160-719-10	Sequence 10, Appl
29	3988	69.0	1075	14	US-10-160-719-34	Sequence 34, Appl
30	3988	69.0	1075	14	US-10-160-719-54	Sequence 54, Appl
31	3963.5	68.6	1081	14	US-10-229-193-6	Sequence 6, Appl
32	3959.5	68.5	1081	14	US-10-229-193-12	Sequence 12, Appl
33	3890	67.3	1052	12	US-10-627-132-30	Sequence 30, Appl
34	3870.5	67.0	1055	16	US-10-437-963-117576	Sequence 117576,
35	3844.5	66.5	1094	12	US-10-627-132-18	Sequence 18, Appl
36	3844.5	66.5	1094	14	US-10-209-059-18	Sequence 18, Appl
37	3844.5	66.5	1094	14	US-10-160-719-26	Sequence 26, Appl
38	3844.5	66.5	1094	14	US-10-160-719-46	Sequence 46, Appl
39	3844.5	66.5	1165	9	US-09-900-237-8	Sequence 8, Appl
40	3830	66.3	847	12	US-10-424-599-238832	Sequence 238832,
41	3820.5	66.1	1039	9	US-09-900-237-14	Sequence 14, Appl
42	3820	66.1	1059	12	US-10-627-132-42	Sequence 42, Appl
43	3820	66.1	1059	14	US-10-209-059-42	Sequence 42, Appl
44	3820	66.1	1059	14	US-10-160-719-2	Sequence 2, Appl
45	3818	66.1	1086	9	US-09-900-237-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-900-237-30
; Sequence 30, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
; LENGTH: 1080
; TYPE: PRT
; ORGANISM: Triticum aestivum
; US-09-900-237-30

Query Match	100.0%	Score 5778;	DB 9;	Length 1080;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1080;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY 1 MDGDADALKSGRHGAGDVQCICADGLGTTLDGDVFTACDVCRFPVCRPCYHERKEGQA 60
|||||
Db 1 MDGDADALKSGRHGAGDVQCICADGLGTTLDGDVFTACDVCRFPVCRPCYHERKEGQA 60
|||||

QY 61 CLOCKTKYKRRHSGPAIRGEEGDDTDADGSDFNYPASGTEDQKQIADRMRSWRMTGG 120
|||||
Db 61 CLOCKTKYKRRHSGPAIRGEEGDDTDADGSDFNYPASGTEDQKQIADRMRSWRMTGG 120
|||||

QY 121 SGNVGHPKVDGSEIGLSKYDSGEIPRGYVPSVTSQMSGEIPGASPDHMMSPGTNISR 180
|||||

Db 121 SGNVGHKPKYDSGEIGLSKYDSGEIPRGVYPSVTNSQMSGEIPGASPDHMHMGPTGNISRR 180
QY 181 APPFYVNHSPNPSRFSGSGIGNVANKERVVDGWMKQDKGAIPWMTNGTSTIAPSEGRAATDI 240
Db 181 APPFYVNHSPNPSRFSGSGIGNVANKERVVDGWMKQDKGAIPWMTNGTSTIAPSEGRAATDI 240
QY 241 DASTEYNMEDALLNDETROPKSRKVPKPIASSKNPYRMWIVLRLVLSIFLHRLTNPNVRN 300
Db 241 DASTEYNMEDALLNDETROPKSRKVPKPIASSKNPYRMWIVLRLVLSIFLHRLTNPNVRN 300
QY 301 AYLWLLSVICETWFAWLSWILDOPPKWFPINRETYLDRALRYDREGEPSQLAAVDIFVS 360
Db 301 AYLWLLSVICETWFAWLSWILDOPPKWFPINRETYLDRALRYDREGEPSQLAAVDIFVS 360
QY 361 TVDPLKEPPIVTANTVLSILAVDYPVDKVSVCYVSDDGASMLTFDALAETSEFARKWVPFV 420
Db 361 TVDPLKEPPIVTANTVLSILAVDYPVDKVSVCYVSDDGASMLTFDALAETSEFARKWVPFV 420
QY 421 KKYDIEPRAPPEFYFCOKIDYLDKQVOPSFKVDRRAMKREYEEFKIRINALVSKALKVPEE 480
Db 421 KKYDIEPRAPPEFYFCOKIDYLDKQVOPSFKVDRRAMKREYEEFKIRINALVSKALKVPEE 480
QY 481 GWIMQDGTWPNGNTRDHPGMIQVFLHSGGLDTEGNEPRLVYVREKRPQFQHHKKAG 540
Db 481 GWIMQDGTWPNGNTRDHPGMIQVFLHSGGLDTEGNEPRLVYVREKRPQFQHHKKAG 540
QY 541 ANNALVRVSAVLNTGQYMLNLDCHYINNSKAVREACFLMDNPLGPOVCYVQPPQFDFG 600
Db 541 ANNALVRVSAVLNTGQYMLNLDCHYINNSKAVREACFLMDNPLGPOVCYVQPPQFDFG 600
QY 601 IDRNDRVANRNTVFFDINLGLDGIQGPVYVGTGCVENRTAIYGYEPPKAKPGFLASL 660
Db 601 IDRNDRVANRNTVFFDINLGLDGIQGPVYVGTGCVENRTAIYGYEPPKAKPGFLASL 660
QY 661 CGGKKKASKSKRSDDKKSKHVDSSVPVFNLEDEEGVEGAGDFDEKSVLMSQMSLEK 720
Db 661 CGGKKKASKSKRSDDKKSKHVDSSVPVFNLEDEEGVEGAGDFDEKSVLMSQMSLEK 720
QY 721 RFGQSAAFVASTLMYEGVGPQSSPESLLKEAIIHVISCGYEDKSEWGTGIIWYIGSVTE 780
Db 721 RFGQSAAFVASTLMYEGVGPQSSPESLLKEAIIHVISCGYEDKSEWGTGIIWYIGSVTE 780
QY 841 YGGRKLKFLERPAYINTTIYPLTSLPLVYCILPAICLLTGKFIPEISNLASIWFLALFL 900
Db 841 YGGRKLKFLERPAYINTTIYPLTSLPLVYCILPAICLLTGKFIPEISNLASIWFLALFL 900
QY 901 SIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHFLFAVFOGLLKVLAGIDTNTFTVSKA 960
Db 901 SIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHFLFAVFOGLLKVLAGIDTNTFTVSKA 960
QY 961 NDEBGDPAELMYKFWKTTLLIPTTILINMVGWVAGTSYAINSGYQSWGPLFGKLFPAFW 1020
Db 961 NDEBGDPAELMYKFWKTTLLIPTTILINMVGWVAGTSYAINSGYQSWGPLFGKLFPAFW 1020
QY 1021 VIVHLYPPLKGLMGQRNTPPTIVVAVLLASIFSLMLWVRVDPPTTRLAGNIIQTCGINC 1080
Db 1021 VIVHLYPPLKGLMGQRNTPPTIVVAVLLASIFSLMLWVRVDPPTTRLAGNIIQTCGINC 1080

RESULT 2
US-10-437-963-174885
; Sequence 174885, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(S3221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 174885
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72784C.1.pep
; US-10-437-963-174885

Query Match 94.4%; Score 5452.5; DB 16; Length 1081;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 999; Conservative 52; Mismatches 29; Indels 1; Gaps 1;

QY 1 MCGDADALKSGRHGAGDVCOICADGLGTTLDGVDFTACDVCRFPVCRPCYEHKKEGTQA 60
Db 1 MCGDADAVKSGRHSGGQACQICGCGVGTTAEGDVFAACDVCGFPVCRPCYEHKKEGTQA 60
QY 61 CLQCKTKYKXHRGSPAIRGEEGDDTDADGSDPNYPASGTEDOKQKIADRMRSWRMTGG 120
Db 61 CPQCKTKYKXHKGPSAIRGEEGDDTDADVDSDNYNPASGSADQKQKIADRMRSWRMNAGG 120
QY 121 SGNVGHKPKYDSGEIGLSKYDSGEIPRGVYPSVTNSQMSGEIPGASPDHMHMGPTGNISRR 180
Db 121 GGDVGRPKYDSGEIGTKYDSGEIPRGYIPSVTNSQISGEIPGASPDHMHMGPTGNIGKR 180
QY 181 APPFYVNHSPNPSRFSGSGIGNVANKERVVDGWMKQDKGAIPWMTNGTSTIAPSEGRAATDI 240
Db 181 APPFYVNHSPNPSRFSGSGIGNVANKERVVDGWMKQDKGAIPWMTNGTSTIAPSEGRGVGDI 240
QY 241 DASTEYNMEDALLNDETROPKSRKVPKPIASSKNPYRMWIVLRLVLSIFLHRLTNPNVRN 300
Db 241 DASTEYNMEDALLNDETROPKSRKVPKPIASSKNPYRMWIVLRLVLSIFLHRLTNPNVRN 300
QY 301 AYLWLLSVICETWFAWLSWILDOPPKWFPINRETYLDRALRYDREGEPSQLAAVDIFVS 360
Db 301 AYLWLLSVICETWFAWLSWILDOPPKWFPINRETYLDRALRYDREGEPSQLAAVDIFVS 360
QY 361 TVDPLKEPPIVTANTVLSILAVDYPVDKVSVCYVSDDGASMLTFDALAETSEFARKWVPFV 420
Db 361 TVDPMKEPPLVTANTVLSILAVDYPVDKVSVCYVSDDGAAMLTFDALAETSEFARKWVPFV 420
QY 421 KKYDIEPRAPPEFYFCOKIDYLDKQVOPSFKVDRRAMKREYEEFKIRINALVSKALKVPEE 480
Db 421 KKYNIETPRAPPEWYFSQKIDYLDKXVHPSFKVDRRAMKREYEEFKIRINGLVAKAQKVEE 480
QY 481 GWIMQDGTWPNGNTRDHPGMIQVFLHSGGLDTEGNEPRLVYVREKRPQFQHHKKAG 540
Db 481 GWIMQDGTWPNGNTRDHPGMIQVFLHSGGLDTEGNEPRLVYVREKRPQFQHHKKAG 540
QY 541 ANNALVRVSAVLNTGQYMLNLDCHYINNSKAVREACFLMDNPLGPOVCYVQPPQFDFG 600
Db 541 ANNALVRVSAVLNTGQYMLNLDCHYINNSKALREACFLMDNPLGRSVCYVQPPQFDFG 600
QY 601 IDRNDRVANRNTVFFDINLGLDGIQGPVYVGTGCVENRTAIYGYEPPKAKPGFLAS 659
Db 601 IDRNDRVANRNTVFFDINLGLDGIQGPVYVGTGCVENRTAIYGYEPPKOKKGSFLUSS 660
QY 660 LCGGKKKASKSKRSDDKKSKHVDSSVPVFNLEDEEGVEGAGDFDEKSVLMSQMSLE 719
Db 661 LCGGKKKASKSKRSDDKKSKHVDSSVPVFNLEDEEGVEGAGDFDEKSVLMSQMSLE 720
QY 720 KRFQSAAFVASTLMYEGVGPQSSPESLLKEAIIHVISCGYEDKSEWGTGIIWYIGSVTE 779
Db 721 KRFQSAAFVASTLMYEGVGPQSSPESLLKEAIIHVISCGYEDKSEWGTGIIWYIGSVTE 780

```

QY 780 DILTGFMHARGWRSVYCMKPAFGSAPINLSDRINQVLRWALGSVEILFSRHCPWY 839
Db 781 DILTGFMHARGWRSVYCMKPAFGSAPINLSDRINQVLRWALGSVEILFSRHCPWY 840
QY 840 GYGRUKFLERFAYINTTIYPLTSLPLLYCILPAICLTGKFMPEISNLSIWFIALF 899
Db 841 GYGRUKFLERFAYINTTIYPLTSLPLLYCILPAICLTGKFMPEISNLSIWFIALF 900
QY 900 LSIFATGILEMRWSGVGIDWWRNEQFWVIGGISAHLEFAVFOGLLKVLAGIDTNFTVTSK 959
Db 901 LSIFATGILEMRWSGVGIDWWRNEQFWVIGGISAHLEFAVFOGLLKVLAGIDTNFTVTSK 960
QY 960 ANDEBGFDAELYMFKWTTLLIPPTTILINMGVWAGTSYAINSGVQSWGPLFGKLFPAF 1019
Db 961 ASDEBGFDAELYMFKWTTLLIPPTTILINMGVWAGTSYAINSGVQSWGPLFGKLFPAF 1020
QY 1020 WIVHLYPFLKGLMGRQNTPTIIVVWVALLASIFSLMWVRDPFTRLAGNPIQTCGIN 1079
Db 1021 WIVHLYPFLKGLMGRQNTPTIIVVWVALLASIFSLMWVRDPFTRTRVGTQTCGIN 1080
QY 1080 C 1080
Db 1081 C 1081

RESULT 3
US-10-627-132-22
; Sequence 22, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1079
; TYPE: PRT
; ORGANISM: Zea mays
US-10-627-132-22

Query Match 93.9%; Score 5423.5; DB 12; Length 1079;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 996; Conservative 46; Mismatches 37; Indels 1; Gaps 1;

QY 1 MDGDADALKSRGHGAGDYCOICADGLGTLDDGVFTACDVCFPVCRPCYHERKBGTQA 60
Db 1 MEGDADGVKSRGGGVQVCGDGVGTAGSDVFTACDVCFPVCRPCYHERKDGTA 60
QY 61 CLQCKTKYKRHRGSPAIRGEEDDTDDADDGSDFNYPASGTDQOKIADMRSMRMTGG 120
Db 61 CPQCKNKYKRHRGSPAIRGEEDDTDDADDGSDFNYPASGTDQOKIADMRSMRMTGG 120
QY 121 SGNVGHPKYDSEGLSKYDSEIPIRGVVPVSNQMSGEIPGASPDHMMSPGNTISRR 180
Db 121 SGVGRPKYDSEGLTKYDSEIPIRGVVPVSNQMSGEIPGASPDHMMSPGNTISRR 180
QY 181 APFVYVNSPNSREFSGISGNVAKERVGDGKWKQDKGATPMNTNGTISIAPSEGRATDI 240
Db 181 APFPMNHSSNPSREFSGVGNVAKERVGDGKWKQDKGTIPMTNGTISIAPSEGRGVGDI 240

```

```

QY 241 DASTEYNNMEDALINDETQPLSRKVPDIASKINPYRMVIVLRVLVLSIFUHYRLTNPVN 300
Db 241 DASTDYNMEDALINDETQPLSRKVPDIASKINPYRMVIVLRVLVLSIFUHYRLTNPVN 300
QY 301 AYPLWLLSVICELWFALSWILDOPFKWFPINRRTYLDRLALRYDREGEPSQLAAVDIFVS 360
Db 301 AYPLWLLSVICELWFALSWILDOPFKWFPINRRTYLDRLALRYDREGEPSQLAAVDIFVS 360
QY 361 TVDPLKEPIVANTVLSILAVDYPVDKSVSCYVSDGASMLTFDALAETSEFAKRWVPFV 420
Db 361 TVDPMKEPEPLVANTVLSILAVDYPVDKSVSCYVSDGASMLTFDALAETSEFAKRWVPFV 420
QY 421 KKYDLEPRAPPEYFCQIDYLDKQVOPSVFKORRAMKREYEEPKIRINALVSKALVPVE 480
Db 421 KKYDLEPRAPPEYFCQIDYLDKQVOPSVFKORRAMKREYEEPKIRINALVSKALVPVE 480
QY 481 GWIMQDGTWPNGNTRDHPGMIQVFLHSGGLDTEGNELPRLVYVYREKRPQGHKKAG 540
Db 481 GWIMQDGTWPNGNTRDHPGMIQVFLHSGGLDTEGNELPRLVYVYREKRPQGHKKAG 540
QY 541 ANNALVRVSAVLTNQGYMLNLDCHYINNSKAVREAMCFMDPNLGPQVYQFPQDFG 600
Db 541 ANNALVRVSAVLTNQGYMLNLDCHYINNSKAVREAMCFMDPNLGPQVYQFPQDFG 600
QY 601 IDNRDRIANRNTVEFDINLRGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFLASL 660
Db 601 IDNRDRIANRNTVEFDINLRGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFLASL 660
QY 661 CGGKKKASKSKRSDDKKSKNHVDSVPVFNLEDEEGVEGAGDFDDEKSLVMSQMSLEK 720
Db 661 CGGKKKASKSKK-GSDKKSKQHVDSSVPVFNLEDEEGVEGAGDFDDEKSLVMSQMSLEK 719
QY 721 RFGQSAAFVASTLMEYGVVQSSPTPELLEKAIHVISCGYEDKSEWGTGIGWYGSVTD 780
Db 721 RFGQSAAFVASTLMEYGVVQSSPTPELLEKAIHVISCGYEDKSEWGTGIGWYGSVTD 779
QY 781 ILTGFMHARGWRSVYCMKPAFGSAPINLSDRINQVLRWALGSVEILFSRHCPWY 840
Db 781 ILTGFMHARGWRSVYCMKPAFGSAPINLSDRINQVLRWALGSVEILFSRHCPWY 839
QY 841 YGGRUKFLERFAYINTTIYPLTSLPLLYCILPAICLTGKFMPEISNLSIWFIALF 900
Db 841 YGGRUKFLERFAYINTTIYPLTSLPLLYCILPAICLTGKFMPEISNLSIWFIALF 899
QY 901 SIFATGILEMRWSGVGIDWWRNEQFWVIGGISAHLEFAVFOGLLKVLAGIDTNFTVTSKA 960
Db 901 SIFATGILEMRWSGVGIDWWRNEQFWVIGGISAHLEFAVFOGLLKVLAGIDTNFTVTSKA 959
QY 961 NDEBGFDAELYMFKWTTLLIPPTTILINMGVWAGTSYAINSGVQSWGPLFGKLFPAF 1020
Db 961 NDEBGFDAELYMFKWTTLLIPPTTILINMGVWAGTSYAINSGVQSWGPLFGKLFPAF 1019
QY 1021 VIVHLYPFLKGLMGRQNTPTIIVVWVALLASIFSLMWVRDPFTRLAGNPIQTCGINC 1080
Db 1021 VIVHLYPFLKGLMGRQNTPTIIVVWVALLASIFSLMWVRDPFTRLAGNPIQTCGINC 1079

RESULT 4
US-10-267-459-6
; Sequence 6, Application US/10267459
; Publication No. US20030150014A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Helentjaris, Timothy G.
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: US/10/267,459
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: US 09/371,383
; PRIOR FILING DATE: 1999-08-06

```

```
; PRIOR APPLICATION NUMBER: US 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1079
; TYPE: PRT
; ORGANISM: Zea mays
US-10-267-459-6

Query Match      93.9%; Score 5423.5; DB 14; Length 1079;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 99%; Conservative 46; Mismatches 37; Indels 1; Gaps 1;

QY 1 MDGDADALSKGRHAGDVCOICADGLGTTLDGVFTACDVCRFPVCRPCYEHKRGTOA 60
Db 1 MEGDADGVKSGRRGGQVCOICGDDGVTGTAEGDVTACDVCGFPVCRPCYEHKRGTOA 60
QY 61 CLOCKTKYKXHRGSPAIRGEGDDTDADDSDENYFASGTEKOKIADRMRSWRNVTGG 120
Db 61 CPQCKNKYKXHRGSPAIRGEGDDTDADDSDENYFASGDDQOKIADRMRSWRNVTGG 120
QY 121 SGNVGHKPYDGSIGLKYDSGEIPRGYVPSVTNSQMSGIPGASPDHMMSPGTNISR 180
Db 121 SGDVGRPKYDSIGLTKYDSGEIPRGYIPSVINSQISGEIPGASPDHMMSPGTNIGR 180
QY 181 APPFYVNHSPNPSREFSGSITGNVAVKRVGKWKQKKGAIPTMTNGTSIAPSEGRA 240
Db 181 APPFYVNHSPNPSREFSGSITGNVAVKRVGKWKQKKGAIPTMTNGTSIAPSEGR 240
QY 241 DASTEYNMEDALLNDETROPILSRKVPATSKINPYRMVILRLVLSIFLHYLTNPVR 300
Db 241 DASTDYNMEDALLNDETROPILSRKVPATSKINPYRMVILRLVLSIFLHYLTNPVR 300
QY 301 AYPILWLLSVICIEIWFALSWILDQPPKWPINRETYLDRLALRYDREGEPSQLAA 360
Db 301 AYPILWLLSVICIEIWFALSWILDQPPKWPINRETYLDRLALRYDREGEPSQLA 360
QY 361 TVDPLKEPPIYVANTVLSILAVDYPVKVSCYVSDGASMLTFDALAETSEFARKW 420
Db 361 TVDPMKEPPLVANTVLSILAVDYPVKVSCYVSDGGAAMLTFDALAETSEFARKW 420
QY 421 KKYDIEPRAPEFYFCQKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALV 480
Db 421 KXNIEPRAPEWYFSQKIDYLDKQVHPSFVKDRAMKREYEEFKIRUNGLVAKAQ 480
QY 481 GWINQDGTTPWGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVSRKRPFGQHK 540
Db 481 GWINQDGTTPWGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVSRKRPFGQHK 540
QY 541 ANNALVRVSAVLITNGQYMLNLDCHYINNSKAVREACFLMDPNLGPQVCYVQ 600
Db 541 ANNALVRVSAVLITNGQYMLNLDCHYINNSKALREACFLMDPNLGRSVYVQ 600
QY 601 IDRNDRYANRNTVFDFINRLGLDGTQGPVYVGTGCVFNRTAIYGEPPKAKG 660
Db 601 IDRNDRYANRNTVFDFINRLGLDGTQGPVYVGTGCVFNRTAIYGEPPKAKG 660
QY 661 CGGKKKASKSKRSDDKKSKNKHVDSSVPVFNLEDIEEGVEGAGFDDKSVLM 720
Db 661 CGGRKKGSKK-K-GSDKKKSQKHVDSVVPVFNLEDIEEGVEGAGFDDKSVLM 720
QY 721 RFGQSAAPVASTLMEYGVPOSSPESLLEKAIHVISCYEDKSEWGTGIEWG 780
Db 720 RFGQSAAPVASTLMEYGVPOSSPESLLEKAIHVISCYEDKSEWGTGIEWG 780
QY 781 ILTGFKHARWRSYVCMKPKPAFKGAPINLSDRNLQVLRWALGSVEILFSRHC 840
Db 780 ILTGFKHARWRSYVCMKPKPAFKGAPINLSDRNLQVLRWALGSVEILFSRHC 840
QY 841 YGGRKFLERFAYINTTYPYTSPLVLYCILPAICLLTGKFIEMPEINLASI 900
Db 840 YGGRKFLERFAYINTTYPYTSPLVLYCILPAICLLTGKFIEMPEINLASI 900

; Query Match      93.9%; Score 5423.5; DB 14; Length 1079;
; Best Local Similarity 92.2%; Pred. No. 0;
; Matches 99%; Conservative 46; Mismatches 37; Indels 1; Gaps 1;
;
; QY 1 MDGDADALSKGRHAGDVCOICADGLGTTLDGVFTACDVCRFPVCRPCYEHKRGTOA 60
; Db 1 MEGDADGVKSGRRGGQVCOICGDDGVTGTAEGDVTACDVCGFPVCRPCYEHKRGTOA 60
; QY 61 CLOCKTKYKXHRGSPAIRGEGDDTDADDSDENYFASGTEKOKIADRMRSWRNVTGG 120
; Db 61 CPQCKNKYKXHRGSPAIRGEGDDTDADDSDENYFASGDDQOKIADRMRSWRNVTGG 120
; QY 121 SGNVGHKPYDGSIGLKYDSGEIPRGYVPSVTNSQMSGIPGASPDHMMSPGTNISR 180
; Db 121 SGDVGRPKYDSIGLTKYDSGEIPRGYIPSVINSQISGEIPGASPDHMMSPGTNIGR 180
; QY 181 APPFYVNHSPNPSREFSGSITGNVAVKRVGKWKQKKGAIPTMTNGTSIAPSEGRA 240
; Db 181 APPFYVNHSPNPSREFSGSITGNVAVKRVGKWKQKKGAIPTMTNGTSIAPSEGR 240
; QY 241 DASTEYNMEDALLNDETROPILSRKVPATSKINPYRMVILRLVLSIFLHYLTNPVR 300
; Db 241 DASTDYNMEDALLNDETROPILSRKVPATSKINPYRMVILRLVLSIFLHYLTNPVR 300
; QY 301 AYPILWLLSVICIEIWFALSWILDQPPKWPINRETYLDRLALRYDREGEPSQLAA 360
; Db 301 AYPILWLLSVICIEIWFALSWILDQPPKWPINRETYLDRLALRYDREGEPSQLA 360
; QY 361 TVDPLKEPPIYVANTVLSILAVDYPVKVSCYVSDGASMLTFDALAETSEFARKW 420
; Db 361 TVDPMKEPPLVANTVLSILAVDYPVKVSCYVSDGGAAMLTFDALAETSEFARKW 420
; QY 421 KKYDIEPRAPEFYFCQKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALV 480
; Db 421 KXNIEPRAPEWYFSQKIDYLDKQVHPSFVKDRAMKREYEEFKIRUNGLVAKAQ 480
; QY 481 GWINQDGTTPWGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVSRKRPFGQHK 540
; Db 481 GWINQDGTTPWGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVSRKRPFGQHK 540
; QY 541 ANNALVRVSAVLITNGQYMLNLDCHYINNSKAVREACFLMDPNLGPQVCYVQ 600
; Db 541 ANNALVRVSAVLITNGQYMLNLDCHYINNSKALREACFLMDPNLGRSVYVQ 600
; QY 601 IDRNDRYANRNTVFDFINRLGLDGTQGPVYVGTGCVFNRTAIYGEPPKAKG 660
; Db 601 IDRNDRYANRNTVFDFINRLGLDGTQGPVYVGTGCVFNRTAIYGEPPKAKG 660
; QY 661 CGGKKKASKSKRSDDKKSKNKHVDSSVPVFNLEDIEEGVEGAGFDDKSVLM 720
; Db 661 CGGRKKGSKK-K-GSDKKKSQKHVDSVVPVFNLEDIEEGVEGAGFDDKSVLM 720
; QY 721 RFGQSAAPVASTLMEYGVPOSSPESLLEKAIHVISCYEDKSEWGTGIEWG 780
; Db 720 RFGQSAAPVASTLMEYGVPOSSPESLLEKAIHVISCYEDKSEWGTGIEWG 780
; QY 781 ILTGFKHARWRSYVCMKPKPAFKGAPINLSDRNLQVLRWALGSVEILFSRHC 840
; Db 780 ILTGFKHARWRSYVCMKPKPAFKGAPINLSDRNLQVLRWALGSVEILFSRHC 840
; QY 841 YGGRKFLERFAYINTTYPYTSPLVLYCILPAICLLTGKFIEMPEINLASI 900
; Db 840 YGGRKFLERFAYINTTYPYTSPLVLYCILPAICLLTGKFIEMPEINLASI 900

; RESULT 5
; US-10-209-059-22
; Sequence 22, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1079
; TYPE: PRT
; ORGANISM: Zea mays
US-10-209-059-22

Query Match      93.9%; Score 5423.5; DB 14; Length 1079;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 99%; Conservative 46; Mismatches 37; Indels 1; Gaps 1;

QY 1 MDGDADALSKGRHAGDVCOICADGLGTTLDGVFTACDVCRFPVCRPCYEHKRGTOA 60
Db 1 MEGDADGVKSGRRGGQVCOICGDDGVTGTAEGDVTACDVCGFPVCRPCYEHKRGTOA 60
QY 61 CLOCKTKYKXHRGSPAIRGEGDDTDADDSDENYFASGTEKOKIADRMRSWRNVTGG 120
Db 61 CPQCKNKYKXHRGSPAIRGEGDDTDADDSDENYFASGDDQOKIADRMRSWRNVTGG 120
QY 121 SGNVGHKPYDGSIGLKYDSGEIPRGYVPSVTNSQMSGIPGASPDHMMSPGTNISR 180
Db 121 SGDVGRPKYDSIGLTKYDSGEIPRGYIPSVINSQISGEIPGASPDHMMSPGTNIGR 180
QY 181 APPFYVNHSPNPSREFSGSITGNVAVKRVGKWKQKKGAIPTMTNGTSIAPSEGRA 240
Db 181 APPFYVNHSPNPSREFSGSITGNVAVKRVGKWKQKKGAIPTMTNGTSIAPSEGR 240
QY 241 DASTEYNMEDALLNDETROPILSRKVPATSKINPYRMVILRLVLSIFLHYLTNPVR 300
Db 241 DASTDYNMEDALLNDETROPILSRKVPATSKINPYRMVILRLVLSIFLHYLTNPVR 300
QY 301 AYPILWLLSVICIEIWFALSWILDQPPKWPINRETYLDRLALRYDREGEPSQLAA 360
Db 301 AYPILWLLSVICIEIWFALSWILDQPPKWPINRETYLDRLALRYDREGEPSQLA 360
QY 361 TVDPLKEPPIYVANTVLSILAVDYPVKVSCYVSDGASMLTFDALAETSEFARKW 420
Db 361 TVDPMKEPPLVANTVLSILAVDYPVKVSCYVSDGGAAMLTFDALAETSEFARKW 420
QY 421 KKYDIEPRAPEFYFCQKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALV 480
Db 421 KXNIEPRAPEWYFSQKIDYLDKQVHPSFVKDRAMKREYEEFKIRUNGLVAKAQ 480
QY 481 GWINQDGTTPWGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVSRKRPFGQHK 540
Db 481 GWINQDGTTPWGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVSRKRPFGQHK 540
QY 541 ANNALVRVSAVLITNGQYMLNLDCHYINNSKAVREACFLMDPNLGPQVCYVQ 600
Db 541 ANNALVRVSAVLITNGQYMLNLDCHYINNSKALREACFLMDPNLGRSVYVQ 600
QY 601 IDRNDRYANRNTVFDFINRLGLDGTQGPVYVGTGCVFNRTAIYGEPPKAKG 660
Db 601 IDRNDRYANRNTVFDFINRLGLDGTQGPVYVGTGCVFNRTAIYGEPPKAKG 660
QY 661 CGGKKKASKSKRSDDKKSKNKHVDSSVPVFNLEDIEEGVEGAGFDDKSVLM 720
Db 661 CGGRKKGSKK-K-GSDKKKSQKHVDSVVPVFNLEDIEEGVEGAGFDDKSVLM 720
QY 721 RFGQSAAPVASTLMEYGVPOSSPESLLEKAIHVISCYEDKSEWGTGIEWG 780
Db 720 RFGQSAAPVASTLMEYGVPOSSPESLLEKAIHVISCYEDKSEWGTGIEWG 780
QY 781 ILTGFKHARWRSYVCMKPKPAFKGAPINLSDRNLQVLRWALGSVEILFSRHC 840
Db 780 ILTGFKHARWRSYVCMKPKPAFKGAPINLSDRNLQVLRWALGSVEILFSRHC 840
QY 841 YGGRKFLERFAYINTTYPYTSPLVLYCILPAICLLTGKFIEMPEINLASI 900
Db 840 YGGRKFLERFAYINTTYPYTSPLVLYCILPAICLLTGKFIEMPEINLASI 900
```

```
Db 421 KYNIEPRAPWFQKIDYLDKDVHPSFVKDRAMKREVEEFPKIRVGLVAKAQKVPPE 480
QY 481 GWIMODGTWPNGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVREKRPQFQHHKAG 540
Db 481 GWIMODGTWPNGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVREKRPQFQHHKAG 540
QY 541 AMNALVRVSAVLNQTGYMLNLDCHYINNSKAVREAMCFMDNPLGPOVCYQFQRFQDG 600
Db 541 AMNALVRVSAVLNQTGYMLNLDCHYINNSKAUREAMCFMDNPLGRSVCYQFQRFQDG 600
QY 601 IDNRDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYBPPIKAKKPGFLASL 660
Db 601 IDNRDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTALYGYBPPIKQKGFLLSL 660
QY 661 CGGKKKASKSKRSKSKKSNKHVDSSVPVFNLEIEEGVEGAGFDDKSVLMSQMSLEK 720
Db 661 CGGRKKGSKSKK-GSDKKKSQKHVDSSVPVFNLEIEEGVEGAGFDDKSKLLMSQMSLEK 719
QY 721 RFGQSAFVASTLMEYGGVQSQSTPESLLKEAHVISCYVEDKSEWGTGIEGTVSVTE 780
Db 720 RFGQSAFVASTLMEYGGVQSQSTPESLLKEAHVISCYVEDKIEWGTGIEGTVSVTE 779
QY 781 ILTGFMHARGMRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCHPLWYG 840
Db 780 ILTGFMHARGMRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCHPLWYG 839
QY 841 YGRLKFLERFAYINTTIYPLTSLPLLYCYILPAICLLTGKTFIMPISNLASIFIALFL 900
Db 840 YGRLKFLERFAYINTTIYPLTSLPLLYCYILPAICLLTGKTFIMPISNFAISIFISLFI 899
QY 901 SIFATGILEMRWSGVGIDEMWNEQFWLGGISAHLFAVFOGLLKVLAGIDTNFTVTSKA 960
Db 900 SIFATGILEMRWSGVGIDEMWNEQFWLGGISAHLFAVFOGLLKVLAGIDTNFTVTSKA 959
QY 961 NDEGDFAELYMFKWTTLLIPPTTILINMVGVAGTSYAINSGYQSWGLFGLKFFFAFW 1020
Db 960 SDEGDFAELYMFKWTTLLIPPTTILINLVGVAGISYAINSGYQSWGLFGLKFFFAFW 1019
QY 1021 VIVHLYPFLKGLMGRQNRTPPTIVVAVLLASIFSLLVWRVDPFTRLAGNIQTCGINC 1080
Db 1020 VIVHLYPFLKGLMGRQNRTPPTIVVAVLLASIFSLLVWRVDPFTRLAGNIQTCGINC 1079
```

RESULT 6

```
; Sequence 10, Application US/10627132
; Publication No. US2004006876A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Haiyin
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1077
; TYPE: PRT
; ORGANISM: Zea mays
US-10-627-132-10
```

Query Match

93.8%; Score 5421.5; DB 12; Length 1077;

```
Best Local Similarity 92.4%; Pred. No. 0;
Matches 998; Conservative 46; Mismatches 33; Indels 3; Gaps 2;

QY 1 MGDADALSKSRHAGDVCOICADGLTGTLDGDFVFTACDVCRFPVPCVCEHREKGTQ 60
Db 1 MEGDADGVKSGRRGGGQVCQICGDGVGTTAEGDVFAACDVCGFPVCEPCYEYERKGTQ 60
QY 61 CLQCKTKYKRRHSGSPAIRGEEGDDTDADGSGFNPYASGTEDQKQIADMRMSRWNVTG 120
Db 61 CPQCKTKYKRRHSGSPAIRGEEGDDTDAD--SDFNYLASGNEQKQIADMRMSRWNVTG 118
QY 121 SGNVGHPKVDSEIGLSKYDSGEIPRGVYPSVTNSQMSGEIPCASPDHMHMSPTGNISRR 180
Db 119 SGVGRPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGEIPCASPDHMHMSPTGNIGRR 178
QY 181 APFPYVNHSPNPSRFSGSGIGNVANKERVDGWMKODKGAIPTWNTGTSIAPSEGRAATDI 240
Db 179 APFPYVNHSPNPSRFSGSGIGNVANKERVDGWMKODKGAIPTWNTGTSIAPSEGRGVGDI 238
QY 241 DASTEYNNMEDALLNDETROPILSRKVPIASSKINPYRMVIVLRLVLSIFLHYLNTNPNR 300
Db 239 DASTDYNNMEDALLNDETROPILSRKVPLPSSRINPYRMVIVLRLVLSIFLHYRITNPNR 298
QY 301 AYPLMLLSVICIWFALSWMILDQFPKWPFPINRETYLDRALBYDREGEQSOLAAVDIFVS 360
Db 299 AYPLMLLSVICIWFALSWMILDQFPKWPFPINRETYLDRALBYDREGEQSOLAAVDIFVS 358
QY 361 TVDPLKEPPIVTANTVLSILAVDYPVKVSCVSDGASMLTFDALAETSEFARKWVPFV 420
Db 359 TVDPMKEPPIVTANTVLSILAVDYPVKVSCVSDGASMLTFDALAETSEFARKWVPFV 418
QY 421 KKYDIEPRAPPEYFCQIDYLDKQVQPSFVKDRAMKREYEEFKIRINLVALSKALKVPEE 480
Db 419 KKYDIEPRAPPEYFCQIDYLDKQVHPSFVKDRAMKREYEEFKVRVNLVAKAQKVPPE 478
QY 481 GWIMODGTWPNGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVREKRPQFQHHKAG 540
Db 479 GWIMODGTWPNGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVREKRPQFQHHKAG 538
QY 541 AMNALVRVSAVLNQTGYMLNLDCHYINNSKAVREAMCFMDNPLGPOVCYQFQRFQDG 600
Db 539 AMNALVRVSAVLNQTGYMLNLDCHYINNSKALREAMCFMDNPLGRSVCYQFQRFQDG 598
QY 601 IDNRDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYBPPIKAKKPGFLASL 660
Db 599 IDNRDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTALYGYBPPIKQKGFLLSL 658
QY 661 CGGKKKASKSKRSKSKKSNKHVDSSVPVFNLEIEEGVEGAGFDDKSVLMSQMSLEK 720
Db 659 CGGRKKKASKSKK-GSDKKKSQKHVDSSVPVFNLEIEEGVEGAGFDDKSKLLMSQMSLEK 717
QY 721 RFGQSAFVASTLMEYGGVQSQSTPESLLKEAHVISCYVEDKSEWGTGIEGTVSVTE 780
Db 718 RFGQSAFVASTLMEYGGVQSQSTPESLLKEAHVISCYVEDKTEWGTGIEGTVSVTE 777
QY 781 ILTGFMHARGMRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCHPLWYG 840
Db 778 ILTGFMHARGMRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCHPLWYG 837
QY 841 YGRLKFLERFAYINTTIYPLTSLPLLYCYILPAICLLTGKTFIMPISNLASIFIALFL 900
Db 838 YGRLKFLERFAYINTTIYPLTSLPLLYCYILPAICLLTGKTFIMPISNFAISIFISLFI 897
QY 901 SIFATGILEMRWSGVGIDEMWNEQFWVIGGISAHLFVAVFOGLLKVLAGIDTNFTVTSKA 960
Db 898 SIFATGILEMRWSGVGIDEMWNEQFWVIGGISAHLFVAVFOGLLKVLAGIDTNFTVTSKA 957
QY 961 NDEGDFAELYMFKWTTLLIPPTTILINMVGVAGTSYAINSGYQSWGLFGLKFFFAFW 1020
Db 958 SDEGDFAELYMFKWTTLLIPPTTILINLVGVAGISYAINSGYQSWGLFGLKFFFAFW 1017
QY 1021 VIVHLYPFLKGLMGRQNRTPPTIVVAVLLASIFSLLVWRVDPFTRLAGNIQTCGINC 1080
Db 1021 VIVHLYPFLKGLMGRQNRTPPTIVVAVLLASIFSLLVWRVDPFTRLAGNIQTCGINC 1080
```

Db 1018 VIVHLYPFLKGLMGRQNRPTTIWVWAILLASIFSLWVRIDPFTTRVTGPDQTCTGINC 1077

RESULT 7

US-10-209-059-10
; Sequence 10, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Dhugge, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; PRIOR FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1077
; TYPE: PR1
; ORGANISM: Zea mays
US-10-209-059-10

Query Match 93.8%; Score 5421.5; DB 14; Length 1077;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 998; Conservative 46; Mismatches 33; Indels 3; Gaps 2;

QY 1 MGDGADALSGRGAGDGVCOICADGLGTTLDGVFTACDVCRFPVCRPCYEHERKEGTQA 60
Db 1 MEGDADGVKSRRGGGQVCOICGDGVGTTAEGDVFAACDVCGFPVCRPCYEHERKEGTQA 60
QY 61 CLQCKTKYKRRHRSQPAIRGEGDDTDADDGSDFNYPASGTEQKQKIADRMRSWRMNTGG 120
Db 61 CPQCKTKYKRRHRSQPAIRGEGDDTDAD--SDFNYLASGNEQKQKIADRMRSWRMNVGG 118
QY 121 SGNVGHPKYGDSIGLSKYDSGEIPRGVPSVNTSQMSGEIPGASPDHMMSPGTGNSRR 180
Db 119 SGDVGRPKYDSIGLTKYDSGEIPRGVPSVNTSQMSGEIPGASPDHMMSPGTGNGKR 178
QY 181 APPYVNHSPNPSREFSGSNGVAKWVGVGKWKQDKGALPMTNGTSIAPSEGRGAATDI 240
Db 179 APPYVNHSPNPSREFSGSNGVAKWVGVGKWKQDKGALPMTNGTSIAPSEGRGVGDI 238
QY 241 DASTEYNMEDALLNDETROPLSRKVPITASSKINPYRMVIVLRLVLSIFLHYRLTNFVN 300
Db 239 DASTDYNMEDALLNDETROPLSRKVPITASSKINPYRMVIVLRLVLSIFLHYRLTNFVN 298
QY 301 AYLWLLSVTCEIWFALSWILDOPPKWFPINRRTYLDRLALYDRGEPGSQLAAVDIFVS 360
Db 299 AYLWLLSVTCEIWFALSWILDOPPKWFPINRRTYLDRLALYDRGEPGSQLAAVDIFVS 358
QY 361 TVDPLKEPPIVNTANTVLSILAVDYPVDKVCYVSDGASMTDPALETSEFAKWKVPFV 420
Db 359 TVDPMKEPPIVNTANTVLSILAVDYPVDKVCYVSDGASMTDPALETSEFAKWKVPFV 418
QY 421 KYVDIEPRAPEFYPCQKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALKVPEE 480
Db 419 KYNIEPRAPEWYFSQKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALKVPEE 478
QY 481 GWINQDGTWPNNTRDPGMIQVFLHSGGLDTEGNEHLPRLVTVSREKRPFGHKKAG 540
Db 479 GWINQDGTWPNNTRDPGMIQVFLHSGGLDTEGNEHLPRLVTVSREKRPFGHKKAG 538
QY 541 ANNALVRVAVLTNGQVWMLNDCDHYTNNKAVREACFLMDPNLGPQVCVQFPQRPFDG 600
Db 539 ANNALVRVAVLTNGQVWMLNDCDHYTNNKAVREACFLMDPNLGRSVVCVQFPQRPFDG 598

QY 601 IDNRDRYANRNTVFFDINIRGLDIOGPVYVGTGCVFNRTAIYGYEPPKAKKPGFLASL 660
Db 599 IDNRDRYANRNTVFFDINIRGLDIOGPVYVGTGCVFNRTAIYGYEPPKAKKPGFLSSL 658
QY 661 CGGKKKASKKSSDKKSNKHVDSVFNLEDEEGVEGAGFDEKSVLMSQMSLEK 720
Db 659 CGGRKKASKKX--GSDKKKSQKHVDSVFNLEDEEGVEGAGFDEKSVLMSQMSLEK 717
QY 721 RFGQSAAFVASTIMGYGVPOQSSTPESLLKEAHHVISCYEDKSEWGTGIGWYGVSTED 780
Db 718 RFGQSAAFVASTIMGYGVPOQSSTPESLLKEAHHVISCYEDKTEWGTGIGWYGVSTED 777
QY 781 ILTGFMHARGWRSVCMKRPAPKGSABINISDRNLQVLRWALGSVEILFSSHCPWLWYG 840
Db 778 ILTGFMHARGWRSVCMKRPAPKGSABINISDRNLQVLRWALGSVEILFSSHCPWLWYG 837
QY 841 YGGRKLFLERFAVINTTIYPLTSLPLVYCIILPAICLLTGKFTMPRIISNLASITWFIALFL 900
Db 838 YGGRKLFLERFAVINTTIYPLTSLPLVYCIILPAICLLTGKFTIPISNFAISWFIISLFI 897
QY 901 SIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHLFVAFQGLLKVLGIDTNTFTVTSKA 960
Db 898 SIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHLFVAFQGLLKVLGIDTNTFTVTSKA 957
QY 961 NDEEGDPAELYMPKWTLLIPPTTILINMVGVVAGTSVAINSGYOSWGLFCKLPFAFW 1020
Db 958 SDEGDGPAELYMPKWTLLIPPTTILINLVGVVAGISVAINSGYOSWGLFCKLPFAFW 1017
QY 1021 VIVHLYPFLKGLMGRQNRPTTIWVWAILLASIFSLWVRIDPFTTRVTGPDQTCTGINC 1080
Db 1018 VIVHLYPFLKGLMGRQNRPTTIWVWAILLASIFSLWVRIDPFTTRVTGPDQTCTGINC 1077

RESULT 8

US-10-160-719-6
; Sequence 6, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugge
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1077
; TYPE: PR1
; ORGANISM: Zea mays
US-10-160-719-6

Query Match 93.8%; Score 5421.5; DB 14; Length 1077;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 998; Conservative 46; Mismatches 33; Indels 3; Gaps 2;

QY 1 MGDGADALSGRGAGDGVCOICADGLGTTLDGVFTACDVCRFPVCRPCYEHERKEGTQA 60
Db 1 MEGDADGVKSRRGGGQVCOICGDGVGTTAEGDVFAACDVCGFPVCRPCYEHERKEGTQA 60
QY 61 CLQCKTKYKRRHRSQPAIRGEGDDTDADDGSDFNYPASGTEQKQKIADRMRSWRMNTGG 120
Db 61 CPQCKTKYKRRHRSQPAIRGEGDDTDAD--SDFNYLASGNEQKQKIADRMRSWRMNVGG 118
QY 121 SGNVGHPKYGDSIGLSKYDSGEIPRGVPSVNTSQMSGEIPGASPDHMMSPGTGNSRR 180
Db 119 SGDVGRPKYDSIGLTKYDSGEIPRGVPSVNTSQMSGEIPGASPDHMMSPGTGNGKR 178

```
Db 119 SGDVRPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGEIPGASPDHMHMPTGNIKR 178
QY 181 APPPYVNHSPNSRFSFGSIGNVAMKERVDMKQKDGKAIPTMTNGTSTAPSEGRAATDI 240
Db 179 APPPYVNHSPNSRFSFGSIGNVAMKERVDMKQKDGKGTIPTMTNGTSTAPSEGRGVGDI 238
QY 241 DASTEYNMEDALLNDETROPLSRKVP IASKNIPYRMVILVRLVLSIFLHVLRLTNPVN 300
Db 239 DASTDYNMEDALLNDETROPLSRKVP LPSRNIPYRMVILVRLVLSIFLHVLRLTNPVN 298
QY 301 AYPMLLSVICIWFALSWILDQPKWFPINRETYLDRALRYDREGESQIAAVIDFVS 360
Db 299 AYPMLLSVICIWFALSWILDQPKWFPINRETYLDRALRYDREGESQIAAVIDFVS 358
QY 361 TVDPLKEPPIVANTVLSILAVDPVKVSCVSDGASMLTFDALAETSEFARKWVPV 420
Db 359 TVDPMKEPPLVANTVLSILAVDPVKVSCVSDGAAWMLTFDALAETSEFARKWVPV 418
QY 421 KKYDIEPRAPEFYFCOKIDYLDKVPQSFVKDRRAMKREYEEFKIRINALVSKALKVPEE 480
Db 419 KKYDIEPRAPEFYFSQIDYLDKVPQSFVKDRRAMKREYEEFKVRVNGLVAKAQKVPPE 478
QY 481 GWIMQDGTWPNGNTRDHPGMIQVFLHSGGGLDTGEGNELPRLVYVSREKRFQHHKAG 540
Db 479 GWIMQDGTWPNGNTRDHPGMIQVFLHSGGGLDTGEGNELPRLVYVSREKRFQHHKAG 538
QY 541 ANNALVRYSAVLTNGQYMLNLDCHYINNSKAVREAMCFMDPNLGPQVCYQFPQRDQ 600
Db 539 ANNALVRYSAVLTNGQYMLNLDCHYINNSKALREAMCFMDPNLGRSVCYQFPQRDQ 598
QY 601 IDRNDRYANRNTVFDDINLRGLDGIQGPVYVGTGVFNRTAIYGEYEPPIKAKKPGFLASL 660
Db 599 IDRNDRYANRNTVFDDINLRGLDGIQGPVYVGTGVFNRTAIYGEYEPPIKQKGFLLSL 658
QY 661 CGKXKASKSKRSDDKKSNKHVDSSVPVFNLEDEIEGVEGAGDDKSVLMSQMSLEK 720
Db 659 CGGRKKASKSKK--GSDKKSKQKHVDSSVPVFNLEDEIEGVEGAGDDKSKLMSQMSLEK 717
QY 721 RFGQSAAFVASTLMYEGVGPQSSTPESLLKEAIIHVISCGYEDKSWGTGTEIGWYGSVTD 780
Db 718 RFGQSAAFVASTLMYEGVGPQSATPESLLKEAIIHVISCGYEDKSWGTGTEIGWYGSVTD 777
QY 781 ILTGFKMHARGWRSVYCMKPKPAFKGSAPINLSDRNLQVLRWALGSVEILFSRHCPLMYG 840
Db 778 ILTGFKMHARGWRSVYCMKPKPAFKGSAPINLSDRNLQVLRWALGSVEILFSRHCPLMYG 837
QY 841 YGGRKLFLEFAYINTTYPILTSPLLYVCILPAICLLTGKPIIPEISNFASIWFIPLFI 900
Db 838 YGGRKLFLEFAYINTTYPILTSPLLYVCILPAICLLTGKPIIPEISNFASIWFIPLFI 897
QY 901 SIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLFVAVFQGLLAVLAGIDTNFTVTSKA 960
Db 898 SIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLFVAVFQGLLAVLAGIDTNFTVTSKA 957
QY 961 NDEBGDFAELMYFKWTTLLIPPTTILINMVGVWAGTSYAINSGYQSGPLFGKLFPAFW 1020
Db 958 SDEGDFAELMYFKWTTLLIPPTTILINLVGVWAGISYAINSGYQSGPLFGKLFPAFW 1017
QY 1021 VIVHLYPPLKGLMGRQNTPTIIVVAVLLASIFSLLWVRDPTTRLAGNIOFTCGINC 1080
Db 1018 VIVHLYPPLKGLMGRQNTPTIIVVWAILLASIFSLLWVRIDPFTTRVTGPDQTCTGINC 1077
```

RESULT 9

```
US-10-160-719-30
; Sequence 30, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugra
; APPLICANT: Timothy G. Helencjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
```

```
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30
; LENGTH: 1077
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-160-719-30
```

```
Query Match 93.8%; Score 5421.5; DB 14; Length 1077;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 998; Conservative 46; Mismatches 33; Indels 3; Gaps 2;

QY 1 MGDADALAKSGRHGAGDVQCICADGLGTTLDGDVFTACDVCRFPVPCRCYHERKEGTQA 60
Db 1 MEGDADGVKSGRRGGQVCQICGDDGVGTGTAAGDVFACDVCGFPVPCRCYERKDGTA 60
QY 61 CLQCKTKYKRHRGSPAIRGEGDDTDADDGSDPNYPASGTEQKOKIADRMESWMTGG 120
Db 61 CPQCKTKYKRHRGSPAIRGEGDDTDAD--SDFNYLASGNEQKOKIADRMESWMTGG 118
QY 121 SGNVGHPKYDSGEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMHMSPGTGNI 180
Db 119 SGDVRPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGEIPGASPDHMHMSPGTGNI 178
QY 181 APPPYVNHSPNSRFSFGSIGNVAMKERVDMKQKDGKAIPTMTNGTSTAPSEGRAATDI 240
Db 179 APPPYVNHSPNSRFSFGSIGNVAMKERVDMKQKDGKGTIPTMTNGTSTAPSEGRGVGDI 238
QY 241 DASTEYNMEDALLNDETROPLSRKVP IASKNIPYRMVILVRLVLSIFLHVLRLTNPVN 300
Db 239 DASTDYNMEDALLNDETROPLSRKVP LPSRNIPYRMVILVRLVLSIFLHVLRLTNPVN 298
QY 301 AYPMLLSVICIWFALSWILDQPKWFPINRETYLDRALRYDREGESQIAAVIDFVS 360
Db 299 AYPMLLSVICIWFALSWILDQPKWFPINRETYLDRALRYDREGESQIAAVIDFVS 358
QY 361 TVDPLKEPPIVANTVLSILAVDPVKVSCVSDGASMLTFDALAETSEFARKWVPV 420
Db 359 TVDPMKEPPLVANTVLSILAVDPVKVSCVSDGAAWMLTFDALAETSEFARKWVPV 418
QY 421 KKYDIEPRAPEFYFCOKIDYLDKVPQSFVKDRRAMKREYEEFKIRINALVSKALKVPEE 480
Db 419 KKYDIEPRAPEFYFSQIDYLDKVPQSFVKDRRAMKREYEEFKVRVNGLVAKAQKVPPE 478
QY 481 GWIMQDGTWPNGNTRDHPGMIQVFLHSGGGLDTGEGNELPRLVYVSREKRFQHHKAG 540
Db 479 GWIMQDGTWPNGNTRDHPGMIQVFLHSGGGLDTGEGNELPRLVYVSREKRFQHHKAG 538
QY 541 ANNALVRYSAVLTNGQYMLNLDCHYINNSKAVREAMCFMDPNLGPQVCYQFPQRDQ 600
Db 539 ANNALVRYSAVLTNGQYMLNLDCHYINNSKALREAMCFMDPNLGRSVCYQFPQRDQ 598
QY 601 IDRNDRYANRNTVFDDINLRGLDGIQGPVYVGTGVFNRTAIYGEYEPPIKAKKPGFLASL 660
Db 599 IDRNDRYANRNTVFDDINLRGLDGIQGPVYVGTGVFNRTAIYGEYEPPIKQKGFLLSL 658
QY 661 CGKXKASKSKRSDDKKSNKHVDSSVPVFNLEDEIEGVEGAGDDKSVLMSQMSLEK 720
Db 659 CGGRKKASKSKK--GSDKKSKQKHVDSSVPVFNLEDEIEGVEGAGDDKSKLMSQMSLEK 717
QY 721 RFGQSAAFVASTLMYEGVGPQSSTPESLLKEAIIHVISCGYEDKSWGTGTEIGWYGSVTD 780
Db 718 RFGQSAAFVASTLMYEGVGPQSATPESLLKEAIIHVISCGYEDKSWGTGTEIGWYGSVTD 777
QY 781 ILTGFKMHARGWRSVYCMKPKPAFKGSAPINLSDRNLQVLRWALGSVEILFSRHCPLMYG 840
```

```
Db 778 ILTGKMHARGWRSIYCMKPAKGSAPINLSDRNLQVRLWALGSEVILFSRHCPLWYG 837
Qy 841 YGRLKFLERFAYINTIYPLTSPLLYCILLPAICLLTGKFMPIBSINLASIWFIALFL 900
Db 838 YGRLKFLERFAYINTIYPLTSPLLYCILLPAICLLTGKFTIPIBSINFASIFSLFI 897
Qy 901 SIFATGILEMRWSGVGIDSEWRNEQFWIIGGISAHLFAVFGGLKVLKVLGIDTNTFTVTSKA 960
Db 898 SIFATGILEMRWSGVGIDSEWRNEQFWIIGGISAHLFAVFGGLKVLKVLGIDTNTFTVTSKA 957
Qy 961 NDEEGDFAELYMFKWTTLLIPPTTILIINMGVYAGTSYAINSGYOSWGFLFKGLFFAFW 1020
Db 958 SDEGDGFAELYMFKWTTLLIPPTTILIINMGVYAGTSYAINSGYOSWGFLFKGLFFAFW 1017
Qy 1021 VIVHLYPFLKGLMGRQNRPTIIVVWAILLASIFSLWVRIDPFTTRVTGPDQTGGINC 1080
Db 1018 VIVHLYPFLKGLMGRQNRPTIIVVWAILLASIFSLWVRIDPFTTRVTGPDQTGGINC 1077

RESULT 10
US-10-160-719-50
; Sequence 50, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Zea mays
US-10-160-719-50

Query Match 91.1%; Score 5261.5; DB 14; Length 1043;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 970; Conservative 42; Mismatches 31; Indels 3; Gaps 2;

Qy 1 MCGDADALSGRGAGDVCOICADGLGTTLDGDFVFTACDVCRFPCVCRPCYEHERKEGTQA 60
Db 1 MEGDAGVSGRGGGQVCOICDGVGTTAGDVFAACDVGCPVCRPCYEHERKGTQA 60
Qy 61 CLOCKTKYKRRHSGPAIRGEGDDTDADGSDGFNYPASGTEDQKQIADRRMRWRMTGG 120
Db 61 CPQCKTKYKRRHSGPAIRGEGDDTDAD--SDFNLAGSNEQDKQIADRRMRWRMTGG 118
Qy 121 SGNVGHPKYSIGILSKYDSGELPRGYVPSVTNSQMSGIPGASPDHNMWSPGTLNRR 180
Db 119 SGDVGRPKYDSGIGLTKYDSGELPRGYVPSVTNSQISGIPGASPDHNMWSPGTLNRR 178
Qy 181 APPYVNHSPNPREFSGSIGVNAWKERVGVKMKQDKGAIPTWNGTSIAPSEGRATDI 240
Db 179 APPYVNHSPNPREFSGSIGVNAWKERVGVKMKQDKGAIPTWNGTSIAPSEGRGVCDI 238
Qy 241 DASTEYNMEDALLNDETQPLSRKVPFIASSKINPYRMVILRLVLVLSIFILHRYLTNPVRN 300
Db 239 DASTDYNMEDALLNDETQPLSRKVPFIASSKINPYRMVILRLVLVLSIFILHRYLTNPVRN 298
Qy 301 AYLWLLSVCEIWFALSWILDQPKWFFPINRETYLDRALRYDREGEPSQLAAVDIFVS 360
Db 299 AYLWLLSVCEIWFALSWILDQPKWFFPINRETYLDRALRYDREGEPSQLAAVDIFVS 358
```

```
Qy 361 TVDPLKEPPIVANTVLSILAVDYPVDKVSVCYVSDDGASMLTFDALAETSEFARKWVPFV 420
Db 359 TVDPMKEPPLVANTVLSILAVDYPVDKVSVCYVSDDGAAMLTFDALAETSEFARKWVPFV 418
Qy 421 KKYDIEPRAPERYFOKIDYLDKQVQSPVKDORRAMKREYEBKIRINALVSKALKVPEE 480
Db 419 KKYNIERAPAEWYFSQIDYLDKQVHPSFVKDORRAMKREYEEFKVNVGLVAKAQVPEE 478
Qy 481 GWMODGTWPQGNTRDHPGMLQVFLGHSGLDTEGNELPRLYVYVSRKRPQFQHHKAG 540
Db 479 GWMQDGTWPQGNTRDHPGMLQVFLGHSGLDTEGNELPRLYVYVSRKRPQFQHHKAG 538
Qy 541 ANNALVRVSAVLITNGQYMLNLDCHYINNSKAVREAMCFMLDPNLGQVCYVQPPORFDG 600
Db 539 ANNALVRVSAVLITNGQYMLNLDCHYINNSKALREAMCFMLDPNLGSRVYVQPPORFDG 598
Qy 601 IDRNRYANRNVFFDINRLGDIQGPYVVGTCVFNRTAIYGYEPPIKAKKPGFLASL 660
Db 599 IDRNRYANRNVFFDINRLGDIQGPYVVGTCVFNRTAIYGYEPPIKAKKPGFLSSL 658
Qy 661 CGGKKKASKSKKSSDKKSNKHVDSSVPVFNLEDEEGVEGAGFDDKSVLMSQMSLEK 720
Db 659 CGGRKKASKSKK--GSDKKSKQKHVDSSVPVFNLEDEEGVEGAGFDDKSVLMSQMSLEK 717
Qy 721 RFQSSAAFAVASTIMEYGGVPOQSATPESLLKEAIIHVISCYVEDKSEWGTGTYGVTED 780
Db 718 RFQSSAAFAVASTIMEYGGVPOQSATPESLLKEAIIHVISCYVEDKSEWGTGTYGVTED 777
Qy 781 ILTGFKHARGWRSVYCMKPAKGSAPINLSDRNLQVRLWALGSEVILFSRHCPLWYG 840
Db 778 ILTGFKHARGWRSIYCMKPAKGSAPINLSDRNLQVRLWALGSEVILFSRHCPLWYG 837
Qy 841 YGRLKFLERFAYINTIYPLTSPLLYCILLPAICLLTGKFMPIBSINLASIWFIALFL 900
Db 838 YGRLKFLERFAYINTIYPLTSPLLYCILLPAICLLTGKFMPIBSINLASIWFIALFL 897
Qy 901 SIFATGILEMRWSGVGIDSEWRNEQFWIIGGISAHLFAVFGGLKVLKVLGIDTNTFTVTSKA 960
Db 898 SIFATGILEMRWSGVGIDSEWRNEQFWIIGGISAHLFAVFGGLKVLKVLGIDTNTFTVTSKA 957
Qy 961 NDEEGDFAELYMFKWTTLLIPPTTILIINMGVYAGTSYAINSGYOSWGFLFKGLFFAFW 1020
Db 958 SDEGDGFAELYMFKWTTLLIPPTTILIINMGVYAGTSYAINSGYOSWGFLFKGLFFAFW 1017
Qy 1021 VIVHLYPFLKGLMGRQNRPTIIVVW 1046
Db 1018 VIVHLYPFLKGLMGRQNRPTIIVVW 1043
```

```
RESULT 11
US-10-627-132-14
; Sequence 14, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
```

QY	960	ANDEEGDFAELNMFKWTLLIPPTTILIIINWGVVAGTSYAINSGYQSWGPLFGKLF	FAF	101
Db	956	ATDEEGDFAELNMFKWTLLIPPTTILIIINLVGVVAGISYAINSGYQSWGPLFGKLF	FAF	1015
QY	1020	WVIVHLXPFLLKMGCRNRTETIVVAVALLASTFSLWLVRVDPFTTFLAGNPIQTCG	IN	1079
Db	1016	WVIVHLXPFLLKMGCRNRTETIVVAVALLASTFSLWLVRVDPFTTFLAGNPIQTCG	IN	1075
QY	1080	C	1080	
Db	1076	C	1076	
RESULT 12				
US-10-209-059-14				
; Sequence 14, Application US/10209059				
; Publication No. US20030163838A1				
; GENERAL INFORMATION:				
; APPLICANT: Dhugga, Kanwarpal S.				
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses				
; FILE REFERENCE: 0864R2				
; CURRENT APPLICATION NUMBER: US/10/209,059				
; CURRENT FILING DATE: 2002-07-31				
; PRIOR APPLICATION NUMBER: 60/096,822				
; PRIOR FILING DATE: 1998-08-17				
; PRIOR APPLICATION NUMBER: 09/371,383				
; PRIOR FILING DATE: 1999-08-06				
; PRIOR APPLICATION NUMBER: 09/550,483				
; PRIOR FILING DATE: 2000-04-14				
; NUMBER OF SEQ ID NOS: 52				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 14				
; LENGTH: 1076				
; TYPE: PRT				
; ORGANISM: Zea mays				
US-10-209-059-14				
Query Match 89.1%; Score 5149; DB 14; Length 1076;				
Best Local Similarity 87.7%; Pred. No. 0;				
Matches 948; Conservative 69; Mismatches 58; Indels 6; Gaps 6;				
QY	1	MDGDADALKSRHGAGDVCQICADGLGTLLDGDVFTACDVCRCPCYEHKEGTQA	60	
Db	1	MDG-GDATNSGKHVAGVCQICGDBGVTAADGLFTACDVCRCPCYEHKEGTQA	59	
QY	61	CLQCKTKYKRHRGSPAIRGEGDDTDADDGSDFNYPASGTEDQKQIADRMRSWRNWTG	120	
Db	60	CPQCKTKYKRHRGSPVPHGEENEDVDADDVSDNYQASGNQDQKQIAERMLTWTNRSR	119	
QY	121	SGNVGHPKYDSGEICLSKYDSGEIPRGVVPVTSNQSMEIGPCASPDHMMSPGNISSR	180	
Db	120	S-DIGLAKYDSGEIGHGKYDSGEIPRGYIPSLTHSQISGEIPGASPD-HMSPVGNIGRR	177	
QY	181	A-PPFYVNHSPNPSREFSGSNGVAKERVGDGKWKQKGAIPMTNGTISIAPSEGRAATD	239	
Db	178	GHQFPYVNHSPNPSREFSGSLGNVAKERVGDGKWK-DKGAIPTMTNGTISIAPSEGRVAD	236	
QY	240	IDASTEYNMEDALLNDETROPLSRKVPITASSKINPYRMVIVLRVLVLSIFLHYRLTNPVR	299	
Db	237	IDASTDYNMEDALLNDETROPLSRKVPITSSKINPYRMVIVLRVAVLCIFLYRITHPNV	296	
QY	300	NAYPLWLLSVICETWFAISWILDQFPKWFPIINRETYLDRALRYDREGPEPSQLAAVDIFV	359	
Db	297	NAYPLWLLSVICETWFAISWILDQFPKWSPINRETYLDRALRYDREGPEPSQLAPVDIFV	356	
QY	360	STVDPLKEPPIVANTVLSILAVDPVDKVCYVSDDGASMLTFDALAETSEFARKWVPF	419	
Db	357	STVDPKMEPPIVANTVLSILAVDPVDKVCYVSDDGASMLTFDALAETSEFARKWVPF	416	
QY	420	VKKYDIEPRAPEFYFCOKIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSKALKVPE	479	
Db	417	CKKNIEPRAPEFYFAQKIDYLDKQVTSFVKERRAMKREYEEFKVRLNGLVAKAQKPE	476	
QY	480	EGWIMQDTPMGNTRDHPGMIQVFLGHSGGLDTEGNELPLRVVVSREKRPGFQHHKKA	539	
Db	477	EGWIMQDTPMGNTRDHPGMIQVFLGHSGGLDVEGNELPLRVVVSREKRPGFQHHKKA	536	
QY	540	GAMNALVRVSAVLITNGQYMLNLDCHYINNSKAVREAMCFLMDPNLGPQVCYVQFPQFD	599	
Db	537	GAMNALVRVSAVLITNGQYMLNLDCHYINNSKALREAMCFLMDPNLGRNVYVQFPQFD	596	
QY	600	GIDRNDRYANRNTVPFDINLRGLDGIQGPVYVGTGCVFNRTAIYGEPIKAKKPGFLAS	659	
Db	597	GIDRNDRYANRNTVPFDINLRGLDGIQGPVYVGTGCVFNRTAIYGEPPVKKKPGFFSS	656	
QY	660	LCGGKKKASKKRSDDKKSNKHVDSSVPFNLEDIEEGVEGAGFDDEKSVLMSQMSLE	719	
Db	657	LCGGRKTSKSKK-SSEKKSHRHADSSVPFNLEDIEEGIEGSQFDDEKSLIMSQMSLE	715	
QY	720	KRFQSAAFVASTLMEYGGVPOSSTPESLLKEAHIVISCGVEDKSEWGTETGWIYGSVTE	779	
Db	716	KRFQSSVFVASTLMEYGGVPOSATPESLLKEAHIVISCGVEDKTDWGTETGWIYGSVTE	775	
QY	780	DILTFGKHARGWSVYCMKRPAPKGSAPINLSRLNQVLRWALGSVEILFSPRHCPLWY	839	
Db	776	DILTFGKHARGWSIYCMKRPAPKGSAPINLSRLNQVLRWALGSVEILFSPRHCPIWY	835	
QY	840	YGGRKLKFLERFAYINTTIYPLTSLPLVYICILPAICLLTGKFINPEISNLSIASIWFIALF	899	
Db	836	YGGRKLKFLERFAYINTTIYPLTSLPLLYCILPAVCLLTGKFIIPKISNLSVWFISILF	895	
QY	900	LSIFATGILEMRWSVGIDEDWRNRQFVWIGGISIAHLFAVFOGLLKVLAGIDTNTVTSK	959	
Db	896	LSIFATGILEMRWSVGIDEDWRNRQFVWIGGISIAHLFAVFOGLLKVLAGIDTSTVTSK	955	

Db 417 CKKYNIEPRAPWYFAQKIDYLDKVKQTSFVKERRAMKREYEFKVRINGLVAKAQKVE 476
QY 480 EGMIMODGTPWPGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVSREKSPGFQHHKA 539
Db 477 EGMIMODGTPWPGNTRDHPGMIQVFLGHSGLDVEGNEPLRLVYVSREKSPGFQHHKA 536
QY 540 GAMNALVRYSAVLTNGQYMLNLDCHYINNSKAVREAMCFMIDPNLGSPQVYQFPQRF 599
Db 537 GAMNALVRYSAVLTNGQYMLNLDCHYINNSKALREAMCFMIDPNLGRNVYQFPQRF 596
QY 600 GIDNRDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPKAKKPGFLAS 659
Db 597 GIDNRDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPKAKKPGFFSS 656
QY 660 LCGGKKKSKSKKSSDKKSNKHVDSSVPFNLEDEBEGVAGAGDDEKSVLMSQMSLE 719
Db 657 LCGGRKKTSSKK-SSEKKKSHRHADSSVPFNLEDEBEGIGSQDDEKSLMSQMSLE 715
QY 720 KREGQSAAFVASTLMEYGGVPQSTPESLLKEAHHVISCYEDKSEWGTGIGWYGSVTE 779
Db 716 KREGQSSVFVASTLMEYGGVPQSATPESLLKEAHHVISCYEDKTDWGTGIGWYGSVTE 775
QY 780 DILTGFMHARGWRSYVCMKRPAPKGSAPINLSDRLNQVLWALGSVEILFSRHCPWY 839
Db 776 DILTGFMHARGWRSYVCMKRPAPKGSAPINLSDRLNQVLWALGSIEILFSRHCPWY 835
QY 840 GYGRLKFLERFAYINTIYPLTSLPLVYCIPLAICLLTGKFMPEISNLASIMFALF 899
Db 836 GYGRLKFLERFAYINTIYPLTSLPLVYCIPLAICLLTGKFMPEISNLASIMFALF 895
QY 900 LSIFATGILEMRWSGVGIDEMWNEQFWIIGGISAHLFAVFGQLLKVLAGIDTNFTVSK 959
Db 896 LSIFATGILEMRWSGVGIDEMWNEQFWIIGGISAHLFAVFGQLLKVLAGIDTSFTVSK 955
QY 960 ANDEEGDFAELYMFKWTTLLIPTTILINMGVWAGTSYAINSGYQSWGLPFGKLPFAF 1019
Db 956 ATDEEGDFAELYMFKWTTLLIPTTILINMGVWAGTSYAINSGYQSWGLPFGKLPFAF 1015
QY 1020 WVIVHLYPFLKGLMGKQNRTPTIIVWAVLLASIFSLWLVRVDPFTTRLAGPNIQTCGIN 1079
Db 1016 WVIVHLYPFLKGLMGKQNRTPTIIVWAVLLASIFSLWLVRVDPFTTRTAGDIACGIN 1075
QY 1080 C 1080
Db 1076 C 1076

RESULT 13

US-10-160-719-58
; Sequence 58, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 1076
; TYPE: PRT
; ORGANISM: Zea mays
US-10-160-719-58

Query Match 89.1%; Score 5149; DB 14; Length 1076;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 948; Conservative 69; Mismatches 58; Indels 6; Gaps 6;
QY 1 MDGADALXSGRHGAGDVCOICADGLGTLTLDGVDFTACDVCRFPVCRPCYEHERKKGTOA 60
Db 1 MDG-GDATNSGKHVAGVCQICGDSGVGTAAADGLFTACDVCGFPVCRPCYEHERKKGTOA 59
QY 61 CLOCKTKYRHRHSGPAIRGBEGDDTDADGSDPNYPASGTEDOKOKIADMRMRWNMTGG 120
Db 60 CPOCKTKYRHRHSGSPVHGEENEDVDADDSDNYOASGNQOKOKIAERMLTWRNRSRG 119
QY 121 SGNVGHPKYDSGEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMMSTGNSRR 180
Db 120 S-DIGLAKYDSGEIGHGKYDSGEIPRGYIPSLTHSQISGEIPGASPD-HMMSVPGNIGRR 177
QY 181 A-PPFYVNSHPNPSRFSGSI GNANKERVDGKMKQDKGAI PMTNGTSTAPSEGRAATD 239
Db 178 GHQFPYVNSHPNPSRFSGSLGNVANKERVDGKMK-DKGAI PMTNGTSTAPSEGRGVAD 236
QY 240 IDASTEYNMEDALLNDETROPISRKYPIASSKINPYRMVIVLRLVLSIFLHYRLTNPVR 299
Db 237 IDASTYNNMEDALLNDETROPISRKYPISSRINPYRMVIVLRLVLCIFLRYITHPVN 296
QY 300 NAYPLMLLSVCEIWFALSMILDOFPKWPFPINRBTYDLRLALRYDREGPSQLAAVDIFV 359
Db 297 NAYPLMLLSVCEIWFALSMILDOFPKWPFPINRBTYDLRLALRYDREGPSQLAPVDIFV 356
QY 360 STVDPLKEPPIVTANTVLSILAIDYVDKVSVCYSDGASMLTFDALAETSEFARKWVPF 419
Db 357 STVDPMKERPLVTANTVLSILAIDYVDKVSVCYSDGASMLTFDALAETSEFARKWVPF 416
QY 420 VKKYDIEPRAPWYFAQKIDYLDKVKQTSFVKERRAMKREYEFKVRINGLVAKAQKVE 479
Db 417 CKKYNIEPRAPWYFAQKIDYLDKVKQTSFVKERRAMKREYEFKVRINGLVAKAQKVE 476
QY 480 EGMIMODGTPWPGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVSREKSPGFQHHKA 539
Db 477 EGMIMODGTPWPGNTRDHPGMIQVFLGHSGLDVEGNEPLRLVYVSREKSPGFQHHKA 536
QY 540 GAMNALVRYSAVLTNGQYMLNLDCHYINNSKAVREAMCFMIDPNLGSPQVYQFPQRF 599
Db 537 GAMNALVRYSAVLTNGQYMLNLDCHYINNSKALREAMCFMIDPNLGRNVYQFPQRF 596
QY 600 GIDNRDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPKAKKPGFLAS 659
Db 597 GIDNRDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPKAKKPGFFSS 656
QY 660 LCGGKKKSKSKKSSDKKSNKHVDSSVPFNLEDEBEGVAGAGDDEKSVLMSQMSLE 719
Db 657 LCGGRKKTSSKK-SSEKKKSHRHADSSVPFNLEDEBEGIGSQDDEKSLMSQMSLE 715
QY 720 KREGQSAAFVASTLMEYGGVPQSTPESLLKEAHHVISCYEDKSEWGTGIGWYGSVTE 779
Db 716 KREGQSSVFVASTLMEYGGVPQSATPESLLKEAHHVISCYEDKTDWGTGIGWYGSVTE 775
QY 780 DILTGFMHARGWRSYVCMKRPAPKGSAPINLSDRLNQVLWALGSVEILFSRHCPWY 839
Db 776 DILTGFMHARGWRSYVCMKRPAPKGSAPINLSDRLNQVLWALGSIEILFSRHCPWY 835
QY 840 GYGRLKFLERFAYINTIYPLTSLPLVYCIPLAICLLTGKFMPEISNLASIMFALF 899
Db 836 GYGRLKFLERFAYINTIYPLTSLPLVYCIPLAICLLTGKFMPEISNLASIMFALF 895
QY 900 LSIFATGILEMRWSGVGIDEMWNEQFWIIGGISAHLFAVFGQLLKVLAGIDTNFTVSK 959
Db 896 LSIFATGILEMRWSGVGIDEMWNEQFWIIGGISAHLFAVFGQLLKVLAGIDTSFTVSK 955
QY 960 ANDEEGDFAELYMFKWTTLLIPTTILINMGVWAGTSYAINSGYQSWGLPFGKLPFAF 1019
Db 956 ATDEEGDFAELYMFKWTTLLIPTTILINMGVWAGTSYAINSGYQSWGLPFGKLPFAF 1015
QY 1020 WVIVHLYPFLKGLMGKQNRTPTIIVWAVLLASIFSLWLVRVDPFTTRLAGPNIQTCGIN 1079

Db 1016 WVIVHLYPFLKGLMGKQNRTPTIIVVWAILLASIFSLMVRIDPFTTIVTGPDIAKCGIN 1075
QY 1080 C 1080
Db 1076 C 1076

RESULT 14

US-10-425-114-58605
; Sequence 58605, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58605
; LENGTH: 1119
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700217152_FLI.pep
US-10-425-114-58605

Query Match 88.8%; Score 5129; DB 12; Length 1119;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 944; Conservative 71; Mismatches 60; Indels 6; Gaps 6;

QY 1 MDGADALSKGRHGAGDVCOICADGLGTTLDGDVFTACDVCRFPVCRPCYHERKEGTQA 60
Db 44 MDG-GDANSNGKHVAGQVCICGQGVGTAAADGLFTACDVCGFPVCRPCYERKGTQA 102
QY 61 CLQCKTKYKRRHSGPAIRGEEDDTDDGSDFNYPASGTEDQKQIADRMESWMTGG 120
Db 103 CPQCKTKYKRRHSGPVPVHGEENEDVDADDVSDYNYQASGQDQKQIAERMLTWTNSRG 162
QY 121 SGNVCHPKYDSGEICLSKYDSGEIPRGVPSVNSQMSGEIPGASPDHMHMSPTGNISRR 180
Db 163 S-DIGLAKYDSGEIGHGYDSGEIPRGVPSLTHSQISGEIPGASPD-HMMSPVGNIGGR 220
QY 181 A-PPYPVNHSPNPSREFSGSIGNVAKERVVDGKWKQDKGAIPMTNGTSIAPSEGRAATD 239
Db 221 GHQFPYVNHSPNPSREFSGSIGNVAKERVVDGKWK -DKGAIPMTNGTSIAPSEGRGVAD 279
QY 240 IDASTEYNMEDALLNDETQPLSRKVPITASSKINPYRMVIVLRVLVLSIFLHYRLTNPVR 299
Db 280 IDASTDYMEDALLNDETQPLSRKVPITSSKINPYRMVIVLRVLVLAFLCFLRYRITHFVN 339
QY 300 NAYPLWLLSVICEIWFALSWILDQPKWFPINRETYLDRLALRYDREGEPSQLAAVDIFV 359
Db 340 NAYPLWLLSVICEIWFALSWILDQPKWSPINRETYLDRLALRYDREGEPSQLAPVDIFV 399
QY 360 STVDPLKEPPIVTANTVLSILAVDPYVDKVCYVSDGASMLTDLALETSEFARKWVPF 419
Db 400 STVDPMKEPPIVTANTVLSILAVDPYVDKVCYVSDGAAMLTDLALETSEFARKWVPF 459
QY 420 VKKYDIEPRAPEFYPCQKIDYLDKQVPSFVKDRRAMKREYEERKIRINALVSKALKVPE 479
Db 460 CKKYNIEPRAPEWYFAQIDYLDKQVTSFVKERRAMKREYEERKVRINGLVAKQVPE 519
QY 480 EGNWMDQCTPWPNNTRDHPGMIQVFLHSGGLDTEGNELPRLVYVSRKRPFGQHHKKA 539
Db 520 EGNWMDQCTPWPNNTRDHPGMIQVFLHSGGLDVEGNELPRLVYVSRKRPFGQHHKKA 579

QY 540 GAMNALVRYSAVLTNGQYMLNDCDHYINNSKAVREAMCFMLDPNLPQVCYVQFPQRD 599
Db 580 GAMNALVRYSAVLTNGQYMLNDCDHYINNSKALREAMCFMLDPNLPGRVNCYVQFPQRD 639
QY 600 GIDRNDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPKAKKPGFLAS 659
Db 640 GIDRNDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPVKKKPGFFSS 699
QY 660 LCGGKKKASKRSKSDSKKSNKVDSSVFPNLEDIEEGVEGAGFDDKSKVMSQMSLE 719
Db 700 LCGGRKTKSKKK-SSEKKESHADSSVPVENLEDIEEGSQFDDKSKVMSQMSLE 758
QY 720 KEPGQSAFVASTLMYEGVGPSSPESLLKEAIIHVISCGYEDKSEWGTGTEIOWIYGSVTE 779
Db 759 KRFQSSVFVASTLMYEGVGPSSPESLLKEAIIHVISCGYEDKTDWGTGTEIOWIYGSVTE 818
QY 780 DILTGFKHARGWRSVYCMKRPAPKGSAPINLSRDLNQLRWALGSEVILFSRHCPWY 839
Db 819 DILTGFKHARGWRSVYCMKRPAPKGSAPINLSRDLNQLRWALGSEVILFSRHCPWY 878
QY 840 GYGBRLKFLERFAYINTIYPLTSLPLVYCIPLPAICLLTGKFIEMPEISNLASIWFIALF 899
Db 879 GYGBRLKFLERFAYINTIYPLTSLPLVYCIPLPAICLLTGKFIEMPEISNLASIWFIALF 938
QY 900 LSIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLPFAVFOGLLKVLADITNTFTVTSK 959
Db 939 LSIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLPFAVFOGLLKVLADITNTFTVTSK 998
QY 960 ANDERGEDFAELYMFKWTTLLIPPTTILINMGVVGAGTSYAINSGVQSGWPLFGKLFPAF 1019
Db 999 ATDESGDFAELYMFKWTTLLIPPTTILINMGVVGAGTSYAINSGVQSGWPLFGKLFPAF 1058
QY 1020 WVIVHLYPFLKGLMGKQNRTPTIIVVWAILLASIFSLMVRIDPFTTIVTGPDIAKCGIN 1079
Db 1059 WVIVHLYPFLKGLMGKQNRTPTIIVVWAILLASIFSLMVRIDPFTTIVTGPDIAKCGIN 1118
QY 1080 C 1080
Db 1119 C 1119

RESULT 15

US-10-437-963-142245
; Sequence 14245, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142245
; LENGTH: 1063
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1063)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4326C.1.pap
US-10-437-963-142245

Query Match 88.7%; Score 5122.5; DB 16; Length 1063;

Best Local Similarity 87.9%; Pred. No. 0;		Matches 943; Conservative 64; Mismatches 51; Indels 15; Gaps 4;	
Qy	9	KSGRGAGVQICADGLGTTLDGVFTACDVCRPVCYEHHERKEGTQACLOCKTKY	68
Db	5	KSGRHGSGQACQILGDGX-----DVGFPVCRPCYERKDGSOACPCQCKTKY	52
Qy	69	KHRGSPAIRGEGDDTDADGSDFNYPASGTTEDQKIADRMRSWRMNTGGSGNVGHPK	128
Db	53	KRHKGSPPLGDESDDVADDDASDVNYPTSGNQDHHKJAERMLTWRMNSGRDDIVHVK	112
Qy	129	YDSGEIGLSKYDSGEIPRGYPVSTNSQMSGIPGASPDHMHMSPGTGNISSRA-PPPYVN	187
Db	113	YDSGEIGHPKYDSGEIPRIYIPLSHSQISGEIPGASPD-HMMSPVGNIARRGHPPPYVN	171
Qy	188	HSPNPSREFSGSIGNVANKERVDGWMQKQKGAIPMTNGTSTAPSEGRAATDIDASTEYN	247
Db	172	HSPNPSREFSGSIGNVANKERVDGWMK-DKGAIPMANGTSTAPSEGRGVGDIDASTDYN	230
Qy	248	MEDALINDETROPLSRKVPITASSKINPYRMVIVLRVLVLSIFLHYRLTNPVRNAYPLWLL	307
Db	231	MEDALINDETROPLSRKVPITASSKINPYRMVIVLRVIVLCIFLHYRLTNPVRNAYPLWLL	290
Qy	308	SVICEIWFALSMILDQPKWFFINRETYLDRLALRYDREGEPSQLAAVDIFVSTVDPLKE	367
Db	291	SVICEIWFALSMILDQPKWSPINRETYLDRLALRYDREGEPSQLAPVDIFVSTVDPKME	350
Qy	368	PIVVTANTVLSILAVDYPDKVSCYVSDGASMLTFDALAETSEFARKVPPVVKYDIEP	427
Db	351	PIVVTANTVLSILAVDYPDKVSCYVSDGAAMLTFDALAETSEFARKVPPVVKYDIEP	410
Qy	428	RAPEFYCOKIDYLDKQVPSFVKDRRAMKREYEEFKIRINALVSKALKVPEEGWIMQDG	487
Db	411	RAPEWYFAKIDYLDKQVQASFVKDRRAMKREYEEFKVRNALVAKAQVPEEGWIMQDG	470
Qy	488	TPWPGNNTRDHPGMIQVFLGHSGGLDTEGNELPRLVYVSREKPPGFQHHKKAGAMNALVR	547
Db	471	TPWPGNNTRDHPGMIQVFLGHSGGLDTEGNELPRLVYVSREKPPGFQHHKKAGAMNALVR	530
Qy	548	VSALVTNGQYMLNDCDHYINNSKAVREACFLMDPNLGPQVCYVQPPQDFGIDRNDRY	607
Db	531	VSALVTNGQYLLNDCDHYINNSKALREACFLMDPNLGRRCYVQPPQDFGIDRNDRY	590
Qy	608	ANRNTVFFDINLGLDGIQGVVGTGCVFNRTAIGYEPPIKAKKPGFLASLGGKKKA	667
Db	591	ANRNTVFFDINLGLDGLQGVVGTGCVFNRTAIGYEPPIKQKRGYFSSLCGGKKKT	650
Qy	668	SKSKRSDKSKSKKHVDSSVPVFNLEDEEGVEGAGFDEKSVLMSQMSLEKRFQOSAA	727
Db	651	KSKSKSTEKKSHKHVDSSVPVFNLEDEEGIEGSGFDDDEKSLMSQMSLEKRFQOSSV	710
Qy	728	FVASTLMYGGVQOSSTPESLLKEAITHVISCYGEDKSEWTEIGWYGVSTEDILTGFKM	787
Db	711	FVASTLMYGGVQOSATPESLLKEAITHVISCYGEDKSDWTEIGWYGVSTEDILTGFKM	770
Qy	788	HARGWRSVCMKRPAPKPGSAPINLSRLNOVLRWALGVSVEILFSRHCPLWYGYGRLKF	847
Db	771	HARGWRSIYCMKRPAPKPGSAPINLSRLNOVLRWALGVSVEILFSRHCPWYGYGRLKF	830
Qy	848	LERFAYINTTIYPLTSLPLVYLCILPAICLLTGKFMPEISNLSIASIWFIALFISIFATGI	907
Db	831	LERFAYINTTIYPLTSLPLLYLCILPAICLLTGKFIIPESISNFSIASIWFISLISIFATGI	890
Qy	908	LEMWRSVGIDDEWRNEQFVIGISAHLEFVQGLLKLVLGIDTNETVTTSKANDERGDF	967
Db	891	LEMWRSVGIDDEWRNEQFVIGISAHLLFAVQGLLKLVLGIDTFTVTTSKASDEERGDF	950
Qy	968	AELYMFKWTTLLIPTTILLINMVGVVAGTSYAINSGVQSGMPLFGKLFPAFWIVHLYP	1027
Db	951	AELYMFKWTTLLIPTTILLINLVGVVAGISYAINSGVQSGMPLFGKLFPAFWIVHLYP	1010
Qy	1028	FLKGLMGRQNTPTTIVVAVLLASIFSLMWVRDPTTTRLAGNIQTCGINC	1080

Db 1011 FLKGLMGRQNTPTTIVVAVLLASIFSLMWVRDPTTTRVTRVGPDTQKCGINC 1063

Search completed: August 23, 2004, 01:17:23
Job time : 97 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 22, 2004, 15:49:02 ; Search time 9999 Seconds
(without alignments)
4681.516 Million cell updates/sec

Title: US-09-900-237A-30
Perfect score: 5778
Sequence: 1 MDGDADALKSGRHGADVQC.....VDPFTTRLAGPNIQTCGNC 1080

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US0900237/runat_18082004_081514_14633/app_query.fasta_l.1223
-DB=GenEmbl -QMT=Fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -IOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0900237 @CGN_1_1_6626@runat_18082004_081514_14633 -NCFU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_hg_hum.*
31: em_hg_inv.*
32: em_hg_other.*
33: em_hg_mus.*
34: em_hg_pln.*
35: em_hg_rtd.*
36: em_hg_mam.*
37: em_hg_vrt.*
38: em_sy.*
39: em_hgo_hum.*
40: em_hgo_mus.*
41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5778	100.0	3626	8	BT009438	Triticum
2	5452.5	94.4	3954	8	AK072356	Oryza sat
3	5434.5	94.1	3264	6	AX653232	Sequence
4	5423.5	93.9	3795	8	AF200533	Zea mays
5	5423.5	93.9	3799	6	AX338680	Sequence
6	5421.5	93.8	3745	8	AF200528	Zea mays
7	5169.5	89.5	3222	6	AX652952	Sequence
8	5165.5	89.4	4282	8	AK069196	Oryza sat
9	5143	89.0	3676	8	AF200529	Zea mays
10	4838	83.7	87792	8	AP003837	Oryza sat
11	4632	80.2	41693	2	AC145384	AC145384 Oryza sat
12	4632	80.2	154555	8	AC135958	AC135958 Oryza sat
13	4577.5	79.2	3723	8	AF150630	Gossypium
14	4530	78.4	3229	8	BT002335	Arabidops
15	4507	78.0	3682	8	AF027174	Arabidops
16	4499	77.9	3614	6	AR267559	Sequence
17	4499	77.9	3614	6	AX030946	Sequence
18	4499	77.9	3614	6	BD022678	Manipulat
19	4458	77.2	3532	8	AY055724	Populus t
20	4089.5	70.8	29292	8	AB018111	Arabidops
21	4008	69.4	3640	8	AY162181	Populus t
22	3993.5	69.1	3725	8	AF200526	Zea mays
23	3992.5	69.1	3732	8	AK099228	Oryza sat
24	3992.5	69.1	3764	8	AK098978	Oryza sat
25	3992.5	69.1	3768	8	AK102140	Oryza sat
26	3992.5	69.1	3802	8	AK067967	Oryza sat
27	3989.5	69.0	3801	8	AK099281	Oryza sat
28	3981.5	68.9	3897	8	AK100188	Oryza sat
29	3981	68.9	3752	8	AF200525	Zea mays
30	3963.5	68.6	3603	6	AR267557	Sequence
31	3963.5	68.6	3603	6	AX030942	Sequence
32	3963.5	68.6	3603	6	BD022676	Manipulat
33	3963.5	68.6	3763	8	BT008654	Arabidops
34	3959.5	68.5	3673	6	AR267560	Sequence
35	3959.5	68.5	3673	6	AX030948	Sequence
36	3959.5	68.5	3673	6	BD022679	Manipulat
37	3946	68.3	3851	6	BD236020	Materials
38	3870.5	67.0	3631	8	AK121170	Oryza sat
39	3844.5	66.5	3812	8	AF200532	Zea mays
40	3840.5	66.3	4029	8	AK121193	Oryza sat
41	3828	66.3	4029	8	AK100877	Oryza sat
42	3823	66.2	3746	8	AK100914	Oryza sat
43	3820	66.1	3538	8	AF200530	Zea mays
44	3818	66.1	3968	8	AF200531	Zea mays
45	3815.5	66.0	4208	8	AK073561	Oryza sat

ALIGNMENTS

BT009438 3626 bp mRNA linear PLN 20-JUN-2003
 LOCUS Triticum aestivum clone wimk4.pk0015.all:fis, full insert mRNA
 DEFINITION sequence.
 ACCESSION BT009438
 VERSION BT009438.1 GI:32128989
 KEYWORDS FLI CDNA.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 REFERENCE 1 (bases 1 to 3626)
 Tingley,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H.,
 Caraher,N.R., Hanafey,M.K. and Hainey,C.F.
 AUTHORS Direct Submission
 TITLE Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and
 JOURNAL Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
 USA

FEATURES
 source Location/Qualifiers
 1..3626
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="wimk4.pk0015.all:fis"

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3626
 Score: 5778.00 Matches: 1080
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0

US-09-900-237A-30 (1-1080) x BT009438 (1-3626)

181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
 600 GCTCCGTTTCCCTATGTGAATCATTCACCAATCCGTCAGGGAGTGTCTCCGCGAGTATT 659
 201 GlyAsnValAlaTrpIysGluArgValAspGlyTrpIysMetLysGlnAspLysGlyVala 220
 660 GGGAAATGTTGCTGGAAAGAGAGTGTGCTGGAAATGAAGCAGGCAAGGTGCG 719
 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
 720 ATTCCCATGACTAATGGGACAGCATGTCTCCCTCTGAAGTCCGCGAGCTACTGACATC 779
 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
 780 GATGCATCTACTGATACACATGGAAGACGCTTTTACTGAATGATGAATCCCGAGCCT 839
 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
 840 CTATCTAGAAAGTCCCATTTGCTTCTCCAAATAAATCCCTACAGAATGGTCATTGTT 899
 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
 900 CTGCGGTGGTGTCTTAGCATCTTCCTGCACCTACCGTCTCACAAATCTGTGCGTAAT 959
 301 AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle 320
 960 GCATACCCACTGTGGCTTTTATCTGTATATGTGAGATTTGGTTGCTTATCTCTGATA 1019
 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
 1020 CTGGATCAGTTCGGAAGTGTGTTCANTCAACCGGAGACCTACCTTGATAGACTGGCT 1079
 341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaValAspIlePheValSer 360
 1080 TTAAGGTATGACCGAGAGGTGAACCGTCTCAGTGGCTGCTGTGACATATTGTGAGT 1139
 361 ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu 380
 1140 ACAGTCGACCCCTTGAGGAGCCACCTATCGTCACTGCCAACACCTGTGCTATCCATCTT 1199
 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyValAspMet 400
 1200 GCTGTTGATTATCCCGTGGACAAGTCTCTTGCTATGTAICTGATGACGAGCTTCAATG 1259
 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
 1260 CTGACTTTTGACGCAATGGCTGAGACTTCAGAGTTTGTAGAAATGGTACCATTTGTG 1319
 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
 1320 AAGAAGTATGACATTGAACCCAGAGCTCCGAGTTTACTTTTCCAGAAATTTGATTAC 1379
 441 LeuLysAspLysValGlnProSerPheValLysAspArgArgAlaMetLysArgGluTyr 460
 1380 CTGAAGACAAAGTCCAGCCTTCATTGTTTAAAGCCCGCGGCCCATGAAGAGAAATAT 1439
 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
 1440 GAAGAAATTAATAATCAGATAAATGCCCTAGTTTCTAAGGCATTGAAAGTCCCGAGAA 1499
 481 GlyTrpIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly 500
 1500 GGATGGATCATGCAAGATGGCACCATGGCCAGGAAACAATACCGAGGATCATCTCTGGA 1559
 501 MetIleGlnValPheLeuGlyHisSerGlyLysLeuAspThrGluGlyAsnGluLeuPro 520
 1560 ATGATTTCAGGTTTCTTGTGTCACAGTGTGGCTTGATCTGAGGGTAAATGAGCTCCCC 1619
 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 540
 1620 CGTTTAGTTTATGTCTCGTGAAGAGCGCTCTGGGTTCCAGCACCAAGAAGAGCTGGT 1679
 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGluTyrMetLeuAsn 560

```

Db      1680  |||||GCAATGAAATCCCTTGTCTGTCTCAGCTGTCTTACTAATGGACATACATGTTGAAT 1739
QY      561   |||||LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
Db      1740  |||||CITGATGTGATCAGTACATCAACAACAGCAAGGCTGTCCGAGAAGCATATGTGCTTCCTA 1799
QY      581   |||||MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
Db      1800  |||||ATGGATCCAAACCTAGTCCGCAAGTCTGTTATGTGCAGTTCCTCCACAAAGGTTTGATGGG 1859
QY      601   |||||IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
Db      1860  |||||ATTGATAGGAATGATCGATATGCAAAACAGAACACACTGTCTTTTGTATTAATCTTGAGG 1919
QY      621   |||||GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
Db      1920  |||||GGCTTTGACGGCAATTCAGACCAAGTTTATGTGGGAAGTGGTGTGTCTTTCACAGAACCA 1979
QY      641   |||||AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
Db      1980  |||||GCTATCTATGGTTATGAGCCCCCAATTAAAGCGAAGACGAGTTCTTTGGCATCACTA 2039
QY      661   |||||CysGlyGlyLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer 680
Db      2040  |||||TGTGGGGCAAGAAGCAAGCAAGTCAAAAGAAAAGGAGCTCAGATAAGAAAAGTCG 2099
QY      681   |||||AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db      2100  |||||AACAAAGCATGTGGACAGTTCTGTTCAGTATTCAATCTCTCAAGACACATAGAGAGGGTGT 2159
QY      701   |||||GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db      2160  |||||GAAAGTGTCTGGTTTGATGATGAGAAATCAGTTCTCATGTCTCAAATGAGCTTAGAAG 2219
QY      721   |||||ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGlnTyrGlyValPro 740
Db      2220  |||||AGATTGTGGCCAGTCAGCAGCAATTTGTGCTCCACTGTGATGAAATATGTTGTTCT 2279
QY      741   |||||GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db      2280  |||||CAGTCTCCACTCCAGAAATCTCTTTTGAAGAAGACTATCCATGTCAATAAGTTGTGCTAT 2339
QY      761   |||||GluAspLysSerGluTrpGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
Db      2340  |||||GAGGCAAGTCTGAATGGGAACTGAGATTGTTGGATCTATGATCTGTCCACAAAGAT 2399
QY      781   |||||IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
Db      2400  |||||ATTCTTACTGGATTCAAGATGCACGAAGAGCTGGCGTTTCAGTCTATTGCAATGCCAAG 2459
QY      801   |||||ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db      2460  |||||CGCCAGCTTTCAAGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTGTG 2519
QY      821   |||||ArgTrpAlaLeuLysSerValGluIleLeuPheSerArgHisCysProLeuTrpTyrGly 840
Db      2520  |||||CGGTGGGCTCTCGTTCGTGTAATCTTTTCAGCCGGCAATGGCCCTTAGGTATGGC 2579
QY      841   |||||TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrIleTyrPro 860
Db      2580  |||||TACGAGGCGCCTCAAGTCTCTCGAGAGATTGCGTTACATCAACACCACTATTACCCA 2639
QY      861   |||||LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db      2640  |||||CTAACCTCTCTCCGCTCTCTAGTCTATTGTATATTGCGCTGTATCTGTCTGCTCACTGGA 2699
QY      881   |||||LysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeu 900
Db      2700  |||||AAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTCAITGGCTCTTCTCT 2759
QY      901   |||||SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrp 920

```

```

Db      2760  TCAATTTTCCCACTCGTATCCTTGAGATGAGGTGGAGTGGTGTGGCATTTGACGAGTGG 2819
QY      921   |||||TrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaValPhe 940
Db      2820  TGCAGGAATGAACAGTCTTGGGTCAATTGGAGGTATCTCTGCACATCTGTTTGGCGTCTTT 2879
QY      941   |||||GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db      2880  CAGGGTCTTCTGAAGGTGCTTGCCTGATCGCAGTATCGACACCAACTTCACTGTCACTCAAAGGCT 2939
QY      961   |||||AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrLeuLeuIle 980
Db      2940  AATGACGAAGAGCGCACCTTGTCTGAGCTCTACATGTTCAAGTGGACGAGCTTCTTCATC 2999
QY      981   |||||ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db      3000  CCTCCGACGACCACTTTTGATCAATTAACATGGTGTGGTGTGGTCTGGCACCTCTCTACGCC 3059
QY      1001  |||||IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrp 1020
Db      3060  ATCAACAGTGGTTACCAATCATGGGGCGCGCTCTTTGGGAAGCTCTCTTTGGCTTCTGG 3119
QY      1021  |||||ValIleValHisLeuTyrProPheLeuLysGlyLeuMetClyArgGlnAsnArgThrPro 1040
Db      3120  GTGATTGTTCACTTATACCATTCTCTCAAGGGTCTTATGGGAGGCTCTCTTTGGCTTCTGG 3179
QY      1041  |||||ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
Db      3180  ACGATTGTCATCGTCTGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3239
QY      1061  |||||ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db      3240  GTTGATTCATTCACTACCGCTCTCGCTGGCCCAAAATATCCAAACCTGTGGCATCAACTGC 3299

```

AK072356 3954 bp mRNA linear PLN 24-JUL-2003
 Oryza sativa (japonica cultivar-group) cDNA clone:J023059102, full insert sequence.

AK072356
 AK072356.1 GI:32982379

FLI CDNA: CAP trapper.

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

1

REFERENCE

AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of
 Agrobiological Sciences Rice Full-Length cDNA Project Team:
 Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
 Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
 Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
 Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group, Otsu,Y., Murakami,K.,
 Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
 Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
 Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
 Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
 Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,
 Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
 Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
 Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Oka,Y.,
 Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
 Yoshino,M. and Hayashizaki,Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from

japonica rice

Science 301 (5631), 376-379 (2003)

22752273

PUBMED

12869764

REFERENCE

AUTHORS

2 (bases 1 to 3954)
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
 Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,

Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Ikeda, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Koda, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nakikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,

Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K. Genomic Sciences Center and Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kouda, M., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

source
1..3954
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023059102"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	3954
Score:	5452.50	Matches:	999
Percent Similarity:	97.22%	Conservative:	52
Best Local Similarity:	92.41%	Mismatches:	29
Query Match:	94.37%	Indels:	1
DB:	8	Gaps:	1

US-09-900-237a-30 (1-1080) x AK072356 (1-3954)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20

437	ATGACGGCGACGCGGATGCCGTGAAGTCGGGAGGCACGGGAGCGGCGAGCGTCCAG	496
21	IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyVspVpPheThrAlaCysAspVal	40
497	ATCTGGCGGCGAGCGGCGTGGGACGACGGCGGAGGGGACGTGTTCCGCCCTCGCAGCTC	556
41	CysArgPheProValCysArgProCysTyrGluHisGluArgGlyGluGlyThrGlnAla	60
557	TGCGGCTTCCCGTGTGCGCGCCCTGCTACGAGTACGAGCGCAAGGATGCACCCAGGCT	616
61	CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu	80
617	TGCCCCCAGTGGCAAGACCAAGTACAAAGCGCCCAAGCGGAGCCCGCGATCCCTGGGAG	676
81	GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr	100
677	GAAGGCGAGTACTGATGCTGATGATGTCAGTCACTACACTACCTCGACTCGCAGT	736
101	GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly	120
737	GCCGACCAAGCAGAAGATTGCTGATAGTACGCGATGGCGCATGAATGCTGGGGT	796
121	SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp	140
797	GGTGGAGACGTCGCGCTCCCAAGTATGACAGTGGCGAGATCGGGCTCACCAGATATGAC	856
141	SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu	160
857	AGTGGCGAGTCCCTCGGGATACATCCCTTCAGTCACTAAATAGCCAGATCTCGGGAGAA	916
161	IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg	180
917	ATCCCTGGAGCTTCCCTCGATCATCATGATGCTCTCTACCGGAAACATTTGGCAACGT	976
181	AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle	200
977	GCTCCATTCCTTATGTGAACCATTCACCAAAATCCATCAAGGGAGTTCTCTGGCAGCAT	1036
201	GlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAla	220
1037	GGAAATGTTGCTGGAAAGAAAGAGTGTGTTGGTGGAACTGAAGCAGGACCAAGGGAGCA	1096
221	IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle	240
1097	ATTCCTCATGACCAACGCGCACAGCATTTGCCCTTCTGAAGTCGGGAGTTGGTGATATC	1156
241	AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro	260
1157	GATGCATCCACTGATTACAATATGGAAGATGCCCTTACTGAATGAAGAACTCCGACGCT	1216
261	LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal	280
1217	CTCTCTAGAAAAGTTCCCTTCCCTTCATCCAGAAATAATCCCTACAGAATGGTCATTGTT	1276
281	LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn	300
1277	CTGGTTTGGTTGTTCTAAGCATTTTCCTTCACTACCGCTATACGAATCCCTGTGCGCAAT	1336
301	AlaTyrProLeuTyrLeuLeuSerValIleCysGluIleTyrPheAlaLeuSerTrpIle	320
1337	CGGTATCGCTCTCGCTTTTATCTGTTATATGAGATTTGGTTGCTTCTCTCGGATA	1396
321	LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla	340
1397	TTGATCAGTTCCTCCGAGTGGTTCCCAATCAACCGTGAACCTACCTTGATAGGTGCA	1456
341	LeuArgTyrAspArgGluGlyProSerGlnLeuAlaAlaValAspIlePheValSer	360
1457	TTGAGGTATGACAGAGAGGTGAGCCATCTCAGTTGGCTGCTGTTGATTTTGTTCAGT	1516
361	ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu	380

Db	1517	ACAGTCGACCCCATGAAGAACTCTCTTTACTGCCAATACCGTGCTGCTCATTTCTT	1576
Qy	381	AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet	400
Db	1577	GCTGTTGATTACCCAGTGACAGAGGTCCTTGCTATGATCTGACGATGGTCTGCAATG	1636
Qy	401	LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal	420
Db	1637	CTGACTTTTGATGTCATTGCTGACAGCTTCAGAGTTTGCTAGAAAGTGGGTTCCATTCGTT	1696
Qy	421	LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr	440
Db	1697	AAGAAGTATAACATTGAGCAAGAGCTCTGAGTGGTACTTCTCCAGAAAATCGAATTAC	1756
Qy	441	LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr	460
Db	1757	TTCAAAGACAAAGTCCACCTTCATTGTTAAAGACCGCTGCGCCATGAAGAGAAATAT	1816
Qy	461	GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu	480
Db	1817	GAAGAATTCAAAGTTAGGATAAATGGCTTGTTGCTNAGGCACAGAAAGTCCCGAGAG	1876
Qy	481	GlyTrpIleMetGluAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly	500
Db	1877	GGATGGATCATGCAAGATGGCACACCATGGCCAGGAAACAAATACTAGGACCATCTCGA	1936
Qy	501	MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro	520
Db	1937	ATGATTGAGTTTTCCTTGGTCAACAGTGGTGCCCTTGATACCGGGTAATGAGCTTCCC	1996
Qy	521	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly	540
Db	1997	CGCCTGGTCTAGTATCTCGTAGAAGCGTCTGGGTTTCAGCACCAAGAGAGCTGGT	2056
Qy	541	AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn	560
Db	2057	GCCATGAATGCTCTTGTGTCGTCTCAGCTGTGCTTACCAATGGACAGTACATGTTGAAT	2116
Qy	561	LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu	580
Db	2117	CTTGATTGTGATCACTACATCAACACAGCAGGCTCTGCGGGAAGCTATGTGCTTCCTT	2176
Qy	581	MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly	600
Db	2177	ATGGATCCAAACCTAGGAAGAGTGCTGTTATGTTTCAGTTCCCAACAAGGTTGATGGT	2236
Qy	601	IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg	620
Db	2237	ATTGATAGGAATGATCGATATCGGAACAGGAACACTGTGTTTTTCGATATTAACCTGAGG	2296
Qy	621	GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr	640
Db	2297	GCTCTTGATGCAATCCAAAGACAGTTTATGTGGGAACCTGGTTGTGTTATCAACAGAACT	2356
Qy	641	AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGly---PheLeuAlaSer	659
Db	2357	GCTCTATATGTTATGATACCCCAATTAAGCAGAGAAGAGGAGGTTCTTGTTCATCA	2416
Qy	660	LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLys	679
Db	2417	CTATGTGGGGCAGGAAGAGGCAAGCAAGTCAAGAAGAAGAGTTCGCAAGAAGAAG	2476
Qy	680	SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGly	699
Db	2477	TCAACACAGACAGTGAGCAGTGCTGTCAGTTTTCATCTGGAATATGAGAGAGGGT	2536
Qy	700	ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu	719
Db	2537	GTTGAAGGTGTGGATTGATGATGAGAAATCACTTCTTATGTTCTCAATGAGCTTGAG	2596
Qy	720	LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyVal	739
Db	2597	AAGAGATTTGGCCAGTCTGCAGCGTTTGTTCCTCCACTCTCATGGAATATGGTGGTGT	2656
Qy	740	ProGlnSerSerThrProGlnSerLeuLysGluAlaIleHisValIleSerCysGly	759
Db	2657	CCTCAATCTGCAACCCCAAGATCTCTTTTGAAGAAGCTATCCATGTGATAAGTTGTGGC	2716
Qy	760	TyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGlu	779
Db	2717	TATGAGGACAAACCAATGGGAACTGAGATTGGGTGATCTACGGTTCGGTGACAGAA	2776
Qy	780	AspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetPro	799
Db	2777	GATATTCTCACTGGATTCAAGATGCATGCGCAGGCTGGAGATCAATCTACTGCGATGCC	2836
Qy	800	LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal	819
Db	2837	AAGCCGCCAGCTTTCAAGGGTCTGCTCCTATCAATCTTTTCAGATCGTCTTAACCAAGTG	2896
Qy	820	LeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyr	839
Db	2897	CTTCGGTGGGCACTTGGTTCTGTTGAAATCTTTTCAGTCGCCATTTGCCCATATGGTAC	2956
Qy	840	GlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr	859
Db	2957	GGCTATGGAGGAGCGCTTAAGTTCTTTGGAGAGATTGGCTACATCAACACCACTATTAT	3016
Qy	860	ProLeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr	879
Db	3017	CCATTGACATCGATCCGCTTCTCATATACTGTGTTTGGCTGCTATCTGTTGCTCACT	3076
Qy	880	GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPhe	899
Db	3077	GGAAAGTTTCATATCCCAAGATTAGCAACTTTGCTAGTATTGTTGTTCTCTCTCTTC	3136
Qy	900	LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGlu	919
Db	3137	ATTTCAAATTTTGGCCACTGGTATCCTTCAGATGAGTGGAGTGGTGGCATCGATGAG	3196
Qy	920	TrpTrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaVal	939
Db	3197	TGGTGGAGAAATGAACAGTCTCTGGGTTATTGGAGGTATATCTGCGCATCTTTTCCGCTC	3256
Qy	940	PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys	959
Db	3257	TTCCAGGGTCTCTCCAGGTGCTTGTGCTGATTCGACACCAATTTCACTGTCACTCAAG	3316
Qy	960	AlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeu	979
Db	3317	GCTTCTGATGAAGATGGCGACTTTTGTGCTGAGCTCTACATGTTTCAAGTGGACACCGCTTCTC	3376
Qy	980	IleProProThrThrIleLeuIleLeuMetValGlyValAlaGlyThrSerTyr	999
Db	3377	ATCCACCGCAGCAGCATCTTGATCATTAACCTGGTTCGGTGTGTTGTTGTTGTTCTATC	3436
Qy	1000	AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe	1019
Db	3437	CGGATCAACAGCGCTACCAATCATGGGACCGCTCTTTGGCAAGCTCTTCTTGGCTTC	3496
Qy	1020	TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr	1039
Db	3497	TGGGTGATTGTCCTTGTATCCCTTCTCAAGGGTCTTATGGGTGGCGCAACCGCACT	3556
Qy	1040	ProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpVal	1059
Db	3557	CCGACCATCGTGTGTTGGCAATCTTCTTGGCTTCGATCTTCTCATTTGCTGTGGGTT	3616
Qy	1060	ArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn	1079
Db	3617	CGCATCATTCATCCACCGGTGTCACCGGCCAGATATACCCCAACATGTGGCATCAAC	3676
Qy	1080	Cys 1080	
Db	3677	TGC 3679	

```

RESULT 3
AX653232          3264 bp      DNA      linear      PAT 22-MAR-2003
LOCUS             Sequence 3102 from Patent WO03000898.
DEFINITION
ACCESSION         AX653232
VERSION           AX653232.1  GI:29156046
KEYWORDS
SOURCE            Oryza sativa
ORGANISM          Oryza sativa
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS           Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
                Katagiri,F., Qian,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
TITLE             Plant genes involved in defense against pathogens
JOURNAL           Patent: WO 0300898-A 3102 03-JAN-2003;
                Syngenta Participations AG (CH)
FEATURES
source            1..3264
                /organism="Oryza sativa"
                /mol_type="unassigned DNA"
                /db_xref="taxon:4530"
ORIGIN
Alignment Scores:
Pred. No.:        0      Length:      3264
Score:            5434.50      Matches:    998
Percent Similarity: 96.60%      Conservative: 52
Best Local Similarity: 91.81%      Mismatches: 30
Query Match:      94.06%      Indels:    7
DB:               6      Gaps:        2

US-09-900-237A-30 (1-1080) x AX653232 (1-3264)
QY      1 MetAspGlyAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
DB      1 ATGAGCGGCGACGGGAGTGGCGTGAAGTGGGAGGACGGGAGCGGAGCGGTGCCAG 60
QY      21 IleCysAlaAspGlyLeuGlyThrLeuAspGlyAspValPheThrAlaCysAspVal 40
DB      61 ATTCGGCGGACGGCGTGGGCGACGACGGCGGAGGGCGACGTGTTCGCGCGCTCGGAGTC 120
QY      41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
DB      121 TGGCGGCTTCGGGTGGCGGCTGCTACGAGTACGAGCGCAAGGATGACCCAGGCT 180
QY      61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
DB      181 TGCCCCAGTGCAGACCAAGTACAGCGCCACAGGGGAGCGCGCGGATCCCGTGGGGAG 240
QY      81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
DB      241 GAAGGCGAGGATACATGATGCTGATGATGTCTAGTGACTACAACTACCCCTGCATCTGGCAGT 300
QY      101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
DB      301 GCCGACAGAGAGAGAGATGTCTGATAGGATGGCGAGTGGCCAGTGAATGCTGGGGGT 360
QY      121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
DB      361 GGTGGAGACGTGGCGGCTCCCAAGTATGACAGTGGCGAGATCGGGCTCACCAAGTATGAC 420
QY      141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
DB      421 AGTGGCGAGATCCCTCGGGGATACATCCCTTCAGTCACTAATAGCCAGATCTCGGGAGAA 480
QY      161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
DB      481 ATCCCTGGAGCTTCCCTGATCATATATGATGCTCTCTACCGGAAACATTTGGCAAGGT 540
QY      181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200

```

```

DB      541 GCTCCATTCCCTATGTGAACCAATTCACCAATTCATCAAGGGAGTTCCTCTGGCAGCAT 600
QY      201 GlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyVala 220
DB      601 GGAAATGTTCCTGGAAAGAAAGAGTTGATGCTGGAAGTCTGAACTGAACGACGACGAGGAGCA 660
QY      221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
DB      661 ATTCCCATGACCAACGGGACAGCATTCGCCCTTCGTGAAGGTCCGGGAGTTGGTGATATC 720
QY      241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGlnThrArgGlnPro 260
DB      721 GATGATCCACATGATTACAAATATGAAGATGCTTACTGAATGATGAAACTGCCAGCCT 780
QY      261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
DB      781 CTCTCTAGAAAGTTCCTCTTCTCATCCAGAAATAAATCCCTACAGATGTGTCATTGTT 840
QY      281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
DB      841 CTGGGTTCGTGTTCTTAAGCATTTTCTTCACTACCGTATTACGAATCCTGTGCGCAAT 900
QY      301 AlaTyrProLeuTyrLeuLeuSerValIleCysGluIleTyrPheAlaLeuSerTrpIle 320
DB      901 GCGTATCCGCTCTGGGCTTTTATCTGTATATGTAGATTTGGTTTGTCTTGTCTCGATA 960
QY      321 LeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
DB      961 TTGGATCAGTTCGGGAGTGGTTTCCAAATCAACCGTGAAACCTACCTTGATAGCTGGCA 1020
QY      341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
DB      1021 TTGAGGTATGACAGAGAGGTGAGCCATCTCAGTTGGCTGCTCTTGACATTTTTCAGT 1080
QY      361 ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu 380
DB      1081 ACAGTCGACCCCAAGGAACCTCTCTTGTGTACTGCCAATACCGTGTGTCATCTCT 1140
QY      381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
DB      1141 GCTGTTGATTACCGAGTGGACAGGCTCTCTTGCTATGTATCTGACGATGCTGTCGAATG 1200
QY      401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTyrValProPheVal 420
DB      1201 CTGACTTTTCATGCAATGGCTGAGACTTCAGAGTTTGTAGAAAGTGGGTTCCTATTCGT 1260
QY      421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
DB      1261 AAGAAGTATAAATTCAGCCAGAGCTCTCGAGTGGTGTCTCTCCAGAAAATCGATAC 1320
QY      441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr 460
DB      1321 TTGAAGACAAAGTCCACCTTCATTTGTTAAAGCCGTCGTGCCATGAAGAGAAATAT 1380
QY      461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
DB      1381 GAAGAAATCAAGTAGGATAAATGGCTTGTGTCTAAGGCACAGAAAGTCCCGAGGAG 1440
QY      481 GlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly 500
DB      1441 GGAATGGATATGCAAGATGGCACACCATGGCCAGGAAACAATACTAGGGACCATCTCTGGA 1500
QY      501 MetIleGlnValPheLeuGlyHisSerGlyLysLeuAspThrGluGlyAsnGluLeuPro 520
DB      1501 ATGATTCAGGTTTTCCTTGTGTACAGTGGTGGCCTTGATACCGAGGGTAAATGAGCTTCCC 1560
QY      521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 540
DB      1561 CGCCTGTGTACGATATCTCGTGAAGCGCTCTGGGTTTTCAGCACCAACAAAGAGGCTGGT 1620
QY      541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
DB      1621 GCCATGAATGCTCTTGTTCGTGCTCAGCTGTGCTTACCAATGGACAGTACATGTTGAAT 1680

```

QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValAlaGluAlaMetCysPheLeu 580
 Db 1691 CTTGATTGTGATCACTATCAACACAGCAGGCTCTCGGGAAGCTATGCTTCCTT 1740
 QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
 Db 1741 ATGGATCCAAACCTAGGAAGAGTGTCTGTTATGTTCAAGTCCCAAGAGGTTCCATGGT 1800
 QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
 Db 1801 ATTGATAGCAATGATCGATATCGAACAGAACTACTGTGTTTTCGATATTAACTTGAGG 1860
 QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
 Db 1861 GGTCTTGATGGCATCAAGACACAGTTTATGTGGAACTGGTGTGTTTCAACAGAACT 1920
 QY 641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGly---PheLeuAlaSer 659
 Db 1921 GCTCTATATGGTTATGAACCCCAATTAAGCAGAGAAAGGAAGTTCTTGTTCATCA 1980
 QY 660 LeuCysGlyGlyLysLysLysAlaSerLysLysLysLysLysArgSerSerAspLysLysLys 679
 Db 1981 CTATGTGGGGCAGGAAGCAAGCAAGTCAAGAAGAGAGAGTTCGGACAAGAAGAG 2040
 QY 680 SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGly 699
 Db 2041 TCAACAAGACAGCTGGACAGTCTGTCCAGTTTCAATCTTGAAGATATAGAGAGGT 2100
 QY 700 ValGluGlyAlaGlyPheAspAspGlyLysSerValLeuMetSerGlnMetSerLeuGlu 719
 Db 2101 GTTGAAGTGTGGATTGATGATGAGAAATCACTTCTTATGCTCAATGAGCTTGAG 2160
 QY 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyVal 739
 Db 2161 AAGAGATTGGCCAGCTGACAGCTTTGTGCTCCACTCTCATGGAATATGGTGGTGT 2220
 QY 740 ProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGly 759
 Db 2221 CTTCAATCTGCAACCCCAAGAACTCTTTTGAAGAAGCTATCCATGTGATAAGTTGTGCG 2280
 QY 760 TyrGluAspLysSerGluTyrGlyThrGlu-----IleGlyTyrIle 773
 Db 2281 TATGAGGACAAGCCGAATGGGAATGAGGCTATGCAACTGCAATGATTGGTGGATC 2340
 QY 774 TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArg 793
 Db 2341 TACGGTTCCGTGACAGAAGATATTCTCACTGGATTCAAGATGCATCGCGAGGCTGGAGA 2400
 QY 794 SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer 813
 Db 2401 TCATCTACTGATGCCCAAGGCCCGCCAGCTTCAAGGGGTCTGCTCTCATCAATCTTCA 2460
 QY 814 AspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArg 833
 Db 2461 GATCGTCTTAACCAAGTGTCTGGTGGGCACCTGGTCTGTGTTGAAATCTTTTCAGTCGC 2520
 QY 834 HisCysProLeuTyrTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyr 853
 Db 2521 CATTTGCCATATGTATGGCTATGAGGACGCCCTTAAGTTCTTGAGAGAGATTGCCCTAC 2580
 QY 854 IleAsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuPro 873
 Db 2581 ATCAACACCATATTATTCATGTAGATCATGATCCCGCTTCTCATATCTGTTTGGCT 2640
 QY 874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle 893
 Db 2641 GCTATCTGTTGCTCACTGGGAAGTTTCATATCCACAGAGATTAGCAACTTGTCTGATGTT 2700
 QY 894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSer 913
 Db 2701 TGGTTTCATCTCTCTCTCATTTCAATTTTGGCACTGGTATCTCTTGAGATGAGGTGGAGT 2760

QY 914 GlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSer 933
 Db 2761 GGTGTGGCATCGATGAGTGGTGGAGGAATGAACAGTTCTTGGGTTATTGGAGGTATATCT 2820
 QY 934 AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn 953
 Db 2821 GGCATCATTTTGGCGTCTTCCAGGCTCTCTCAAGGTGCTTGTCTGGTATCGACACCAAT 2880
 QY 954 PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe 973
 Db 2881 TTCCTGTCACTCAAGGCTTCTGATGAAGATGCGACTTTGCTGAGCTTACATGTTTC 2940
 QY 974 LysTyrThrThrLeuLeuLeuProThrThrIleLeuIleLeuAsnMetValGlyVal 993
 Db 2941 AAGTGGACAACGCTTCTCATCCACGACGACCATCTTGTATCATTTAACTGTGTCGGTGT 3000
 QY 994 ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGly 1013
 Db 3001 GTTGTCTGTATCTCATACGCGATCAACAGCGCTACCATTCATGGGACCGCTCTTTGGC 3060
 QY 1014 LysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMet 1033
 Db 3061 AAGCTCTCTTTGGCTTCTGGGTGATTGTCCACTTGTACCCCTTCTCAAGGCTCTTATG 3120
 QY 1034 GlyArgGlnAsnArgThrProThrIleValIleValTyrAlaValLeuLeuAlaSerIle 1053
 Db 3121 GGTGGCAAAACCGCACCTCCGACCATCGTTGTGTGGCAATCCTTCTGCTTCGATC 3180
 QY 1054 PheSerLeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle 1073
 Db 3181 TTCCTATTGCTGGTTCGCATCGATCCATCCACCCGCTGTCACCGCCCGCAGATACC 3240
 QY 1074 GlnThrCysGlyIleAsnCys 1080
 Db 3241 CAAACATGGCATCAACTGC 3261

RESULT 4
 AF200533
 LOCUS Zea mays cellulose synthase-9 (Cesa-9) mRNA linear PLN 31-AUG-2000
 DEFINITION Zea mays cellulose synthase-9 (Cesa-9) mRNA, complete cds.
 ACCESSION AF200533
 VERSION AF200533.1 GI:9622889
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 3795)
 AUTHORS Holland, N., Holland, D., Helentjaris, T., Dhugga, K.S., Xoonostle-Cazares, B. and Delmer, D.F.
 TITLE A comparative analysis of the plant cellulose synthase (Cesa) gene family
 JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)
 MEDLINE 20398328
 PUBMED 10938350
 REFERENCE 2 (bases 1 to 3795)
 AUTHORS Dhugga, K.S. and Helentjaris, T.G.
 TITLE Direct Submission
 JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
 FEATURES
 Location/Qualifiers
 1..3795
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="taxon:4577"
 /note="identified and isolated from the maize genome database cDNA collection of Pioneer Hi-Bred International, Inc"
 1..3795
 /gene="Cesa-9"
 238..3477
 /gene="Cesa-9"
 gene 1..3795
 CDS 238..3477

```

/note="CESA-9"
/codon_start=1
/product="cellulose synthase-9"
/protein_id="AAF89969.1"
/db_xref="GI:9622890"
/translation="MEGDADGVKSGRRGGQVCQICGDGVGTTAEGDVFTACDVCGFP
VRCPCYVERKQDTAQPCQKNKYKRHKGSPIALRGEEDDTDADSDFNYPASGNDD
QKQKIADRMRSWRMAGSGDVGRPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGE
IPGADPDHMMSPGTGNI GRRAEPPYMNHSSNPSREFSGVNVAKRVDGWMKKODK
GTPMNGHSIAPSEGRGVGDIDASTDYNMEDALLNDETRQPLSRKVPKLPSSRINPR
MWLVLRIVLSIFLHRYIINPVNARYLWLLSVICEIFALSWILDLQFPRWFPINRER
YLDRLALRYDEGEPSQLAAVDIFVTVDMKEPPLVNTANTVLSILAVDPVDKVSFY
VSDGAAMLTDFALASTFEARWVFKYNIIEPAPWYFSQKIDYLDKVKHPFSFV
KORRAARVEEFKIRVNGLVAKAQVPEGIMODGTPWGNNTRDHPGMIQVFLGH
SGLDTEGNEHLPELVVSREKRPFOHKKKAGAMNALVRVSAVLNNGQYMLNLDCHY
INNSKALRAMCFMPLDNLGRSVYVQFPQREIDRNDRYANRNVFEDINLRGLDG
IGQPVTVTGCVNFRALGYEPPIKKKGGFLLSCGGRKKGSKSGSDKRSQKH
VDSVPVFNLEDEEGVEGAFDEKLSMQMSLEKRFQGSAAFAVSTLMEYGVGQ
SATPESLKEAIAHVIISCGYEDKIEWTEIGWISVTEIDILTGFKHARGWRSIYCMP
KRPAPKGSAPINLSDNLQVLRWALGSVEILFSRHCPLMWGYGRKLFKFLERFAYINTT
IYPLTSLPLIYCILPAICLLTGKFTIIPISNPFASITWISLFSIFATGILEMRWGV
GIDEMWNEQFWVIGISALHFAVQGLLKLAVGIDTNETVTSKASDEGDPALRYMF
KMTLLIPTTILLINLVGVVAGISAINSGVQSWGLFKGLFFAFWVIVHLVPFLKG
LMGRQNRTPTIIVVMWAILLASIPSLWVRIDPFTNRVTGPDTRTCGINC"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 0 Length: 3795
Score: 5423.50 Matches: 996
Percent Similarity: 96.48% Conservative: 46
Best Local Similarity: 92.22% Mismatches: 37
Query Match: 93.86% Indels: 1
DB: 8 Gaps: 1

US-09-900-237A-30 (1-1080) x AF200533 (1-3795)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 238 ATGGAGGGCGACGCGGAGCGCGTGAAGTCGGGAGCGCGGGGAGGCGAGGTGTGCCAG 297

QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 298 ATCTGCGCGGATGCGCGTGGGCACTACGCGCGAGGAGAGCGTCTTCACCGCCTCGCAGCT 357

QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGlyThrGlnAla 60
Db 358 TGGCGGTTCCGGTGTGCGCCCTCTCTACGATGACGAGCGGCGAAGCGGACACAGCG 417

QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 418 TGCCCCCAGTGCACAAAACAAAGTACAAAGCGCCACAAAGGGAGTCCAGCGATCCGAGGGAG 477

QY 81 GluGlyAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 478 GAAGGAGACGATGATGCGGATGATGCTAGGAGCTTCAACTACCGCTGCATCTGGCAAT 537

QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
Db 538 GACGACCAAGACAGAGATTGCTGACAGGATCGCGAGCTGGCGCATGATGCTGGGGGC 597

QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 598 AGCGGGGATGTTGGCGCGCCCAAGTATGACAGTGGTGGATCGGGCTTACCAAGTACGAC 657

QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 658 AGTGGTGAGATCCCTCGGGATACATCCCGTCAGTCAGTACAGCGGCTTACCAAGTACGAC 717

QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 718 ATCCCTGGTGTCTCCCTGACCATCATATGATGTCTCTACTTGGGACACATTGGCAGCGC 777

QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200

```

```

Db 778 GCCCATTTCCCTATATGAATCATTCATAATCCCTCAGGGAATTCCTCGTAGCGTT 837
QY 201 GlyAsnValAlaThrLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAla 220
Db 838 GGGAAATGTTCCCTCGGAAAGAGAGGGTTGATGGCTGAAAATGAAGCAGGACCAAGGAACA 897
QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
Db 898 ATTCCCATGACGAAATGGCACAGCATTTGCTCCTCTGAGGGCGGGGTGTGGTGATATT 957
QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
Db 958 GATGCATCAACTGATTACCAATGGAAGATGCTTATTAAACCATGAACTGCCAGCCT 1017
QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
Db 1018 CTATCTAGGAAAGTTCCACTTCCTCCTCAGGATAATCCATACAGGATGGTCATTGTG 1077
QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
Db 1078 CTACGATTGATTGTTCTAAGCATCTTCTTGCACTACCGGATCACAAATCCTGTGCGTAAT 1137
QY 301 AlaTyrProLeuThrLeuLeuSerValIleCysGluIleThrPheAlaLeuSerTyrIle 320
Db 1138 GCATACCCACTGTGGCTTCTATCTGTTATATGTAGATCTGTTGCTCTTCTCTCGATA 1197
QY 321 LeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
Db 1198 TTGGATCAGTTTCCAAAGTGGTTTCCAAATCAACCGGAGACTTACCTTGATGACTGCCA 1257
QY 341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
Db 1258 TTAAGGTATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGCTGTGACATTTTGTGAGT 1317
QY 361 ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu 380
Db 1318 ACTGTGCGACCAATGAAGGAGCCTCCTCTTGTCTCAGTCCCAATACCGTGTATCCATTCTC 1377
QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
Db 1378 GCTGTGGACTATCCTGTGGATAGGTCTCTTGCTATGTATCTGATGATGAGCTGCTGATG 1437
QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTyrValProPheVal 420
Db 1438 CTGACATTTGATGCACTAGCTAGACTTCAGAGTTTGTCTAGAAAATGGGTGCCATTGTT 1497
QY 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
Db 1498 AAGAAGTACAACTTGAACCTAGAGCTCCTGAAATGGTACTTCTCCAGAAAAATTGATTAC 1557
QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr 460
Db 1558 TTGAAGACAAAGTGCACCTTCATTTGTTTAAAGCCCGCGGCCATGAAGAGAGATAT 1617
QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
Db 1618 GAAGAATTCAAATTTAGGGTAAATGGCTTGTGTGCTTAAGGCACAAAAAGTCCCTGAGAA 1677
QY 481 GlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly 500
Db 1678 GGATGATCATGCAAGATGGCACCATGCGCCAGGAAACAATACCGAGGACCATCCTCGGA 1737
QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
Db 1738 ATGATTTCAGTTTCTTCTGTCAGTGGTGGTCTTGATACTGAGGGTAAATGAGTACCC 1797
QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 540
Db 1798 CGTTTGGTCTATGTTTCTCGTGAACAACTCCTCGGATTCAGCATCACAAAGAAGCTGGT 1857
QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560

```

Db 1858 GCCATGAATGCTCTTGTCGCGCTCTCAGCTGTGCTTACCAATGCGACAATACATGTTGAAT 1917

Qy 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
|||||
Db 1918 CTGTGATTGATCACTTACATCAACACAGTAAGGCTCTCAGGGAAGCTATGCTTCCTT 1977
Qy 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
1978 ATGGATCTTAACCTAGGAAGAGTGTCTGCTATGTTTCAGTTTCCCAGAGGTTCCATGGT 2037

Qy 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
2038 ATTGATAGAAATGATCATGATATCCCAACAGCAACACCGTGTCTTTTCGATATTAACTTGAGA 2097

Qy 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
2098 GGTCTGTATGGCATCCAGGACAGCTTATGTGGGCACCTGGCTGTGTTTCAACAGAAACA 2157

Qy 641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
2158 GCTCTATATGTTATGACCCCCCAATTAAAGCAAAAGAGGTTGGTTTCTGTATCACTA 2217

Qy 661 CysGlyGlyLysLysAlaSerLysSerLysLysArgSerSerAspLysLysSer 680
2218 TGTGTGGCAGGAAGGAAGCAAAATCAAAAGAG---GGCTCAGACAAAGAAAAGTCA 2274

Qy 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
2275 CAGAAAGCATGTGGACAGTTCTGTGGCAGTATTCAATCTTGAAGATATAGAGAGGAGTT 2334

Qy 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
2335 GAAGCGCTGGATTGTATGATGATGAGAAATCACTTCTTATGTCTCAATGAGCTTGGAGAG 2394

Qy 721 ArgPheGlyGlnSerAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro 740
2395 AGATTGGCCAAATCTGCAGCTTTTGTGCTCCACTCTGATGAAATATGCTGCTTCT 2454

Qy 741 GlnSerSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
2455 CAGTCTGGCAGTCCAGAAATCTCTCTGAAGAAGCTATCCATGTATAGTTGTGGCTAC 2514

Qy 761 GluAspLysSerGluThrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
2515 GAGGACAGATTGATGGGAACCTGAGATTGGTGGATCTATGTTCTGTGACGAAAGAT 2574

Qy 781 IleLeuThrGlyPheLysMethHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
2575 ATTCTCACTGGGTTCAAGATGCACGACGAGGCTGGCGTCTGATCTACTGCAATGCTTAAG 2634

Qy 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
2635 CGCGCGGCTTCAAGGGATCGGCTCCCAATCAATCTCTCAGACCGTCTGAACAGGTGCTC 2694

Qy 821 ArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTrpGly 840
2695 CGGTGGGCTCTCGGTTCAGTGGAATCTTTTCAGCGGCATTTGCCCTATGTTACGGG 2754

Qy 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
2755 TACGGAGGACGCTGAAAGTTCTTGGAGAGATTCGCTACATCAACACCACTCTACCCG 2814

Qy 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
2815 CTCAGTCCCTCCGCTCTCTCAATTACTGTATCTCTGCTGCAATCTGCTGCTCAGGGG 2874

Qy 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleThrPheIleAlaLeuPheLeu 900
2875 AAGTTCAATCATCCAGAGATCAGCAACTTCGCTAGTATCTGTTTCATCTCTCTCAIC 2934

Qy 901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrp 920
2935 TCGATCTTCGCCACGGGTATCTCGAGATGAGGTGGAGCGGCTGGGCATCGACAGTGG 2994

Qy 921 TrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaValPhe 940
2995 TGGAGAAACGAGCAGTTCTGGGTCAATCGAGGCACTCCGCCACCTCTTCGCCCTTCTTC 3054

Qy 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
3055 CAGGGCTCTCTCAAGTGCTTGGCGGCATCGACCACTTCAACGTCACCTCCAGGGC 3114

Qy 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
3115 TCGGATGAAGACGGGCACTTCGGGAGCTGTACATGTTCAAGTGGACGACACTTCTGATC 3174

Qy 981 ProProThrThrIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
3175 CGCCCAACACCATCTCATCAACCTGGTGGCGTGTGTGGCGGCATCTCCACGGC 3234

Qy 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
3235 ATCAACAGGGGTACCACTGCTGGGTCGCTCTTCGGCAAGCTCTTCTTCGCTTCTGG 3294

Qy 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
3295 GTGATCGTTCCTGTACCCGTTCTCAAGGCTCTCATGGGTCTCATGGGTGGCAACCGCACCCG 3354

Qy 1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
3355 ACCATCGGTGTCTGGCGCATCTCTGCGTCGATCTTCTCTGCTGGGTTCGC 3414

Qy 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
3415 ATCGATCGTTCACCAACCGGCTCACTGGCCCGGATCTCGAACGCTGTGGCATCAACTGC 3474

RESULT 5
AX338680
LOCUS
DEFINITION
SEQUENCE 5 from Patent WO0179516.
AX338680
ACCESSION
VERSION
AX338680.1 GI:18129042
KEYWORDS
SOURCE
Zea mays
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
Dhugga, K.S. and Helentjaris, T.G.
Maize cellulose synthases and uses thereof
Patent: WO 0179516-A 5 25-OCT-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
1..3799
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
238..3477
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD20545.1"
/db_xref="GI:18129043"
/db_xref="REMBL:CAD20545"
/translation="MEGDAGVKSRRGGGVQICGQGVGTTAEGDVFTACDVGFP
VCRPCYERKDGKQACPCQCKNKYKHKSGPAIRGEGDDTDADADSFNTPSAGNDD
QKQIADRMRSWRNAGSGDVGRPKYDSIGLTKYDSGEIPRGTIPSPVNSQISGE
IFGASPDHMMSPPTGIRRAPFPYMHSSNPSREFSGVGNVAKERVDGKKWKQDK
GTIPMTNGTSTAPSEGRGVGDI DASTYDNMEDALLNDETROP LSKVKYPLPSRINPYR
MVIVLRILVLSIFLHYITNPVNAVPLWLLSVICELWFSWILDVDPKWPINRET
YLDRLALYDREGPFSQLAANDIFVSVDPKKEPLTANTVLSILAVDYPVKVSCY
VSDGAAMLTFDALAETSEFAKWPVFKVKNIEPFAPEVFSQIDYLFKQVHPSFV
KDRAMKREYEEFKIRVNGLVAKQKQEGWIMQDGTWPFGNTRPHGMIOVFLGH
SGGLDTGNEPLRIVYSEKRPQKHKAGAMNALVRVSATVNTGNGMLNLDCHY
INNSKALREAMCFIAMDNLGRSCYVOPFORFDGIDRNDRVARNTVFFDINLRGLDG
IQGPVYVGTGCVFNRTALYGYEPPIKQKGGFLSLCGGRKKGSKSGSKSKQKH

QY 641 AlalIeTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerIleu 660
 Db 2158 GCTCTATATGTTATAGCCCCCAATTAAGCAAGAAGGTTGTTCTTCTCATCACTA 2217
 QY 661 CysGlyGlyLysLysLysAlaSerLysLysLysLysArgSerSerAspLysLysSer 680
 Db 2218 TGTGGTGCAGAGAGAGAGAGCAATCAAGAGAG---GGCTAGACAGAGAAAGTCA 2274
 QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
 Db 2275 CAGAAGCATGTGGACAGTCTGTGTCAGTATTCAATCTTGAAGATATAGAGAGGAGCTT 2334
 QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
 Db 2335 GAAGGCGTGGATTGTGATGATGAGAAATCACTTCTTATGTCTCAATAGCTTGAGAG 2394
 QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro 740
 Db 2395 AGATTGGCCAACTGGACAGCTTTGTTCGCTCCACTCTGATGGAATATGGTGGTTCCT 2454
 QY 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
 Db 2455 CAGTCTCGGACTCCAGAAATCTCTCTGAAGAAGCTATCCATGTCAATAGTTGTGGTAC 2514
 QY 761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
 Db 2515 GAGGACAAAGATTGAATGGGAACTGAGATTGGGTGGATCTATGGTCTGTGACGGAAGAT 2574
 QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
 Db 2575 ATTCTCATCTGGTTCAAGATGACGACGAGGCTGGCGGTGCGATCTACTGATGCTTAAG 2634
 QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
 Db 2635 CGCGCGGCTTCAAGGATCGGCTCCCATCAATCTCTCAGACCGTCTGAACAGGTGCTC 2694
 QY 821 ArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyrGly 840
 Db 2695 CGGTGGGCTCTCGGTTCAAGTGAATCCCTTTTCAGCGGCAATGCCCCCTATGGTACGG 2754
 QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
 Db 2755 TAGGAGAGACGCTGAGTCTTGGAGAGATTGCGCTACATCAACACCATCTACCCG 2814
 QY 861 LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
 Db 2815 CTCAGTCTCCCTCCGCTCTCATTTACTATCTCTGCTGCCATCTGCGTCTCACGGG 2874
 QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeu 900
 Db 2875 AAGTTTCATCATCCAGAGATCAGCAACTTCGTTAGTATCTGGTTATCTCTCTCTTCATC 2934
 QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrp 920
 Db 2935 TCGATCTTCGACCGGATCTCTGGAGATGAGGTGGAGCGGCGTGGCATCGACGAGTGG 2994
 QY 921 TrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaValPhe 940
 Db 2995 TGGAGAACAGCAGTCTCGGTCTATCGGAGGCACTCCGCGCCACCTCTTCGCGCTTTC 3054
 QY 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAspPheThrValThrSerLysAla 960
 Db 3055 CAGGGGCTCTCAAGGTGCTTGGCGGATCGACACCACTTACCCTGTCACCTCCAGAGCC 3114
 QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
 Db 3115 TCGGATGAGACGCGCACTTCGGGAGCTGTACATGTTCAAGTGGACGACACTTCTGATC 3174
 QY 981 ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerfyrAla 1000
 Db 3175 CCGCCACCACCATCTGATCATCAACCTGGTCGGGTTGTTGCCGGCATCTCTACGCC 3234

QY 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
 Db 3235 ATCAACAGCGGTTACAGTCTGTGGGTCCGCTCTTCGGCAAGCTCTTCTTCGCTTCGG 3294
 QY 1021 VallIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
 Db 3295 GTGATCGTTACCTTACCCCTTCTCAAGGCTCTCATGGGTCTGGCAGAACCGCACCCCG 3354
 QY 1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
 Db 3355 ACCATCGGTGTTCGTGGCGATCTGCTGCGCTGATCTTCTCTCTGCTGGGTTCGC 3414
 QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
 Db 3415 ATCGATCGTTTACCACCGGCTCACTGGCCCGGATATCTCGAACGTGTGGCATCAATGC 3474
 RESULT 6
 AF200528 3745 bp mRNA linear PIN 31-AUG-2000
 LOCUS Zea mays cellulose synthase-4 (Cesa-4) mRNA, complete cds.
 DEFINITION Zea mays cellulose synthase-4 (Cesa-4) mRNA, complete cds.
 ACCESSION AF200528
 VERSION AF200528.1 GI:9622879
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 3745)
 AUTHORS Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S.,
 Xocoostle-Cazares,B. and Delmer,D.P.
 TITLE A comparative analysis of the plant cellulose synthase (Cesa) gene
 family
 JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)
 MEDLINE 20398328
 PUBMED 10938350
 REFERENCE 2 (bases 1 to 3745)
 AUTHORS Dhugga,K.S. and Helentjaris,T.G.
 TITLE Direct Submission
 JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred
 International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
 FEATURES
 source
 1. 3745
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="taxon:4577"
 /note="identified and isolated from the maize genome
 database cDNA collection of Pioneer Hi-Bred International,
 Inc"
 gene
 1. 3745
 /gene="Cesa-4"
 CDS
 320. 3553
 /gene="Cesa-4"
 /note="Cesa-4"
 /codon_start=1
 /product="cellulose synthase-4"
 /protein_id="AAF89964.1"
 /db_xref="GI:9622880"
 /translation="MEGDADGVKSGRRGGQVCQICGQVGTAGDVFAACDVGFP
 VCRPCYERKDGTCQCKTKYKRGSPAIRGEGDDTDADSFNKLASNBQK
 OKIADRMRSWRNVCGGVGRPKYDSGEIGLTKYDSGEIIPRGYTPSVNTSISGEIP
 GASPDHMSPTGNIKRAPPPYVNHSPNPSREFSGSICNVAKRVDGKWKQKGT
 IPMTNGTSTAPSEGGVGDIDASTDNMEDALNDETQPLSRKVPFLSSRNIPYRMV
 IVRLVLISIFLHYRITPNVRYAPLWLLSVICETWFLSWILDDQPKWFPINREYL
 DLRLYDREGEPSQLAVDIPVSTVDEMPKPLPVANTVLSILAVDIPVDKVSCTVS
 DDGAAMLTDALETSEFAKWVPEVKNIEPRAPWFYFSQKIDYLDKVKHPSVCS
 RRAMKREVEFKVRVNGLVAKQVPEEGTMDQGTWPNNTRDHPGMIOVFLHSG
 GLDTGRELPRLVYVSRKPGFQHHKAGAMVALRVYSAVLNQQYMLNDGDHYIN
 NSKALREAMCFLMDNLGRSVCYVQPPQRFQIDNDRYANETVTFEINLRLGLDGIQ
 GEVYVTCVFNRTALYGTPEPKKKGKGLSSLCGGRKKAKSKKSGDKKSKQKSHV
 SSVPYFNLEIEEGVEGAFDDEKSLMSQMSLEKRFQGSAAFAVFTLMYEGVQPSA
 TPESLLKEALHVISCGYEDKTEWTEIGWISGVTEIDLTGFMHARGWRSTYCMFKR
 PAFKGSAPINLSRLNQVLRWALGSVEILFSRHCPLWYGYGRLKFLERFAVINTIY

Db 2234 GCTCTATATGGTTATGAGCCCCCAATTAAAGCAGAGAGGGTGGTTCTTCTTCATCACTA 2233
 QY 661 CysGlyGlyLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer 680
 Db 2294 TGTGGCGTAGGAAGAGCAAGCAATCAAGAAG--GGCTCGGACAAAGAAGTCG 2350
 QY 681 AsnLysHisValAspSerValProValPheAsnLeuGluAspLleGluGluGlyVal 700
 Db 2351 CAGAAGCATGTGGACAGTTCTGTGCCAGTATTCAACTTGAAGATATAGAGGGAGTT 2410
 QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
 Db 2411 GAAGCGCTGGATTGACGACGAGAAATCACTTCTTATGTCTCAATGAGCGCTGAGAG 2470
 QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro 740
 Db 2471 AGATTGGCCAGTCGCCAGCGTTTGTTCCTCCACTCTGTATGGAGATATGGTGGTTCCT 2530
 QY 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaLleHisValLleSerCysGlyTyr 760
 Db 2531 CAGTCCGCACTCCGGAGTCTCTTCTGAAAGAAGCTATCCATGTTATAGCTGTGGCTAT 2590
 QY 761 GluAspLysSerGluTyrGlyThrGluLleGlyTyrLleTyrGlySerValThrGluAsp 780
 Db 2591 CAGGACAGACTGATGGGAACTGAGATCGGTGGATCTACGGTTCGTGACAGAGAC 2650
 QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
 Db 2651 ATTCTACCGGATTCAAGATGCACGCGAGGCTGGCGTTCGATCTACTGTCATGCCCAAG 2710
 QY 801 ArgProAlaPheLysGlySerAlaProLleAsnLeuSerAspArgLeuAsnGlnValLeu 820
 Db 2711 CGGCCAGCTTTCAAGGGGTCTGCCCCCAATCAATCTTCGACCCGTCGACACAGGTGCTC 2770
 QY 821 ArgTyrAlaLeuGlySerValGluLleLeuPheSerArgHisCysProLeuTyrTyrGly 840
 Db 2771 CGGTGGGTCTTGGTTCGTGGAGATCTCTTCAGCCGCACTGCCCCCTGGTGGTACGGC 2830
 QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
 Db 2831 TAGCGAGGCGGCTCAAGTTCCTGGAGAGATTGCGGTATCATCAACACCACTATACCGG 2890
 QY 861 LeuThrSerLeuProLeuLeuValTyrCysLleLeuProAlaLleCysLeuLeuThrGly 880
 Db 2891 CTCACGTCCATCCCGTCTCTCATCTACTGCATCTCCGCCGCCATCTGCTGCTCACCGGA 2950
 QY 881 LysPheIleMetProGluLysSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
 Db 2951 AAGTTTCATCTCCAGAGATCAGCACTTCGCCAGCATCTGGTTTCATCTCCCTCTTCATC 3010
 QY 901 SerIlePheAlaThrGlyLleLeuGluMetArgTyrSerGlyValGlyLleAspGluTyr 920
 Db 3011 TCGATCTTCGCCACCGGCATCTCGAGATGAGTGGAGCGGGGTGGGCATCGACGAGTGG 3070
 QY 921 TrrArgAsnGluGlnPheThrValLleGlyGlyLleSerAlaHisLeuPheAlaValPhe 940
 Db 3071 TGGAGGAACGAGCAGTCTGGGTGATCGGGGATCTCCGCCACCTCTCGCGGTGTC 3130
 QY 941 GlnGlyLeuLeuLysValLeuAlaGlyLleAspThrAsnPheThrValThrSerLysAla 960
 Db 3131 CAGGGCTCTCAAGTGTGGCGCGCATCGACACCACTTCACCGTCACTCCAAAGGC 3190
 QY 961 AsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrrThrThrLeuLeuLle 980
 Db 3191 TCGGACGAGGACGGCAGCTTCGGGAGCTGTACATGTTCAAGTGGACGAGCTCTCTGATC 3250
 QY 981 ProProThrThrIleLeuLleLleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
 Db 3251 CCGCCCAACCACTCTGTGATCAACCTGGTGGGTGTCGCGGTGTCGCCGGCATCTCCTAGCC 3310
 QY 1001 IleAsnSerGlyTyrGlnSerTrrGlyProLeuPheGlyLysLeuPhePheAlaPheTrr 1020
 Db 3311 ATCAACAGCGGATACCACTGTCGGGGCCCGCTCTTTCGGCAAGCTCTTCTTCGCTTCGG 3370

QY 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
 Db 3371 GTCAITGTCACCTGTATCCCGTTCCTCAAGGGCTCTATGGCAGCAGACCGCACCCCG 3430
 QY 1041 ThrIleValIleValTrrAlaValLeuLeuAlaSerIlePheSerLeuLeuTrrValArg 1060
 Db 3431 ACCATCGTGTGCTGCGGCATCTGTGGCGTCCATCTTCTCTGCTGTGGGTTCGC 3490
 QY 1061 ValAspProPheThrArgLeuAlaGlyProAsnIleGlnThrCysGlyLeuAsnCys 1080
 Db 3491 ATCAGCCCTTCCACCCCGCTCATCGGCCCGGATACCCAGACGTGTGGCATCACTGC 3550
 RESULT 7
 LOCUS AX652952 3222 bp DNA linear PAT 22-MAR-2003
 DEFINITION Sequence 2822 from Patent WO03000898.
 ACCESSION AX652952
 VERSION AX652952.1 GI:29155766
 KEYWORDS Oryza sativa
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
 Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
 Plant genes involved in defense against pathogens
 Patent: WO 0300898-A 2822 03-JAN-2003;
 JOURNAL Syngenta Participations AG (CH)
 FEATURES
 Location/Qualifiers
 1..3222
 source /organism="Oryza sativa"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4530"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3222
 Score: 5169.50 Matches: 950
 Percent Similarity: 94.52% Conservative: 67
 Best Local Similarity: 88.23% Mismatches: 54
 Query Match: 89.47% Indels: 5
 DB: 6 Gaps: 4
 US-09-900-237A-30 (1-1080) x AX652952 (1-3222)
 QY 6 AspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIleCysAlaAspGly 25
 Db 4 GACGGTGGCAATCGGGAAGCAA-----TGCCATGCTGCGCAGATCTGCGCGCAGCGC 57
 QY 26 LeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCysArgPheProVal 45
 Db 58 GTGGGCACGCGCGCGACGCGAGCTCTTCACCGCTGCGAGCTGCGGGTTCGCGGTG 117
 QY 46 CysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAlaCysLeuGlnCysLys 65
 Db 118 TCGCGGCCATGCTACAGTACGAGTACGAGCGCAAGCAGCGCAGCGTCCCGCAGTCAAG 177
 QY 66 ThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGluGluGlyAspAspThr 85
 Db 178 ACCAGTACAAGCGCCACAAAGGGAGCGCCCGCATCTTGGGGATGAAAGCATGATGTT 237
 QY 86 AspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThrGluAspGlnLysGln 105
 Db 238 GATCGCGATGATGCTAGTGTGTAAGTATCAACATCCCGCAACCCAGACCATAGAC 297
 QY 106 LysIleAlaAspArgMetArgSerTrrArgMetAsnThrGlyGlySerGlyAsnValGly 125
 Db 298 AAGATTGCCGAGAGGATGCTCCTCGCGCATGAATCTGGGAGGAATGATGATATGTC 357
 QY 126 HisProLysTyrAspSerGlyGluLleGlyLeuSerLysTyrAspSerGlyGluLlePro 145

Db 358 CATTTCTAAGTATGATAGTGGTGCAGATTGGTTCATCCCAAGTATGACAGTGGTGAATCCCT 417
Qy 146 ArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGluIleProGlyAlaSer 165
Db 418 CGCATAATATCCCATCGCTCACTCAAGCCAGATCTCAGGTGAATTCCTGGAGCGTCC 477
Qy 166 ProAspHisHisMetMetSerProThrGlyAsnIleSerArgArgAla---ProPhePro 184
Db 478 CCTGAT---CATATGATGTCCTCCGTTGGGACATTGGCAGACGTGGGCATCCATTTCCC 534
Qy 185 TyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIleGlyAsnValAla 204
Db 535 TATGTGAACCATTCACCAACCCCATCAAGGAGTTCCTGGTGCATGTCATGTTGCA 594
Qy 205 TrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAlaIleProMetThr 224
Db 595 TGGAAAGAGAGAGTGTAGTGGCTGGAAATGAAG---GATAAAGGTGCAATTCCTCATGGCT 651
Qy 225 AsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThr 244
Db 652 AATGGTACTAGCATTCCTCTCAGAAGTCTGGGGTTCGTCGATGTCATCTACT 711
Qy 245 GluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLys 264
Db 712 GACTATAACATGGAGGAGCGCTTACTCAATCATGAACACGCCAACCTCTATCGAGAAA 771
Qy 265 ValProIleAlaSerSerLysIleAsnProTyrArgMetValIleValIleLeuArgLeuVal 284
Db 772 GTGCCAATTCGTCATCCAGAAATAATCCGTACAGGATGTTATTTGGTCCCGATTGATT 831
Qy 285 ValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeu 304
Db 832 GTCCTATGTATATTCCTTGCACTACCGTATCACAAATCCTGTGGTAAATGCATACCGCTA 891
Qy 305 TrpLeuLeuSerValIleCysGluIleTyrPheAlaLeuSerTrpIleLeuAspGlnPhe 324
Db 892 TGGTTCCTCTCTCATATGTAGATTTGGTTTGCTCTATCTCTGGATTCCTGGATCAGTTC 951
Qy 325 ProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAsp 344
Db 952 CCCAAGTGTGCCCATCAACCGTGAACCTTACTCGATAGGTGGCTTTAAGGTATGAC 1011
Qy 345 ArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspPro 364
Db 1012 CGGGAAGGTCAACCATCTCAATTGGCTCCTGTTGACATTTTGTGAGTACTGTGGATCCT 1071
Qy 365 LeuLysGluProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyr 384
Db 1072 ATGAAGAAGACCTCCTCTGTCACTGCCTCAATCTGTCTCTCCATCCTGCTGGATTAC 1131
Qy 385 ProValAspLysValSerCysTyrValSerAspAspGlyValAserMetLeuThrPheAsp 404
Db 1132 CCTGTTGATAAGGTATCTTGCTATGTCTGATGATGGAGCTGCAATGTTAATTTGAT 1191
Qy 405 AlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAsp 424
Db 1192 GCGCTGTGCTGAGACTTCAGAAATTTGTAGAAAATGGGTACCGTTCCTGTAAGGTACAGC 1251
Qy 425 IleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLys 444
Db 1252 ATAGAACCCAGAGCTCCAGAGTGGTACTTTGCACAGAAAATTTGATTACTTTGAAGACAAA 1311
Qy 445 ValGlnProSerPheValLysAspArgAlaMetLysArgGluTyrGluGluPheLys 464
Db 1312 GTTCAGGCTCTTTTGTAAAGATCGTGTGTCATGAAGAGGGAATATGAAGAAATTA 1371
Qy 465 IleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGluGlyTrpIleMet 484
Db 1372 GTTCGTGTAAATGCTCTCGTCCCAAGGCACAGAAATTTCTCGAGAGGATGGATTATG 1431
Qy 485 GlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnVal 504

Db 1432 CAAGATGGCACACCTTGGCCTGGTAAACAATACCAGAGACCATCTCGAATGATTCAGGTT 1491
Qy 505 PheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyr 524
Db 1492 TTCCTTGGTCATAGTGGAGGCTTGATACCAGAGCAATAGCTTCCTCGTTAGTCTAT 1551
Qy 525 ValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAla 544
Db 1552 GTGTCTCGTGAGAAACGCTCTGGATTCCAAACACATAAAGGCTGTGTATGAATGCA 1611
Qy 545 LeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAsp 564
Db 1612 CTGTTCGTGTATCAGCTGCTCTTACTAATGGACAATCTTGTGTAATCTTGACTGTGAT 1671
Qy 565 HisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsn 584
Db 1672 CACTACATCAACAATAGCAAAAGCTCTCCGAGAGGCTATGTGCTTCCTTATGGATCCAAAC 1731
Qy 585 LeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsn 604
Db 1732 CTAGGAAGGCGTGTCTGTATGTCCCAATCCCTCAGAGGTTTGACGGTATCGATAGAAAT 1791
Qy 605 AspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGly 624
Db 1792 GATCGATATGCAAAACAGGAACACCGTGTTCGATATTAATTTGAGAGGCTCTCGATGGT 1851
Qy 625 IleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGly 644
Db 1852 CTCCAAGGACCACTTTATGTGGAACCGGTGTGTGTTCACACAGAACAGCTCTTTATGGT 1911
Qy 645 TyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeuCysGlyGlyLys 664
Db 1912 TATGAACCCCCCATTAAGCAGAGAGCGCAGGTACTCTCTTCGCTTTGTGGGGACGA 1971
Qy 665 LysLysAlaSerLysSerLysArgSerSerAspLysLysLysSerAsnLysHisVal 684
Db 1972 AAGAAGACAAAAAGTCGAAGGAGAGAGACCGAAGAAAGAAAGTCACACAAACATGTG 2031
Qy 685 AspSerSerValProValPheAsnLeuGluAspIleGluGluGlyValGluGlyAlaGly 704
Db 2032 GACAGTTCGTGCCAGTTTAAATCTTGAAGATATAGAGGAGGATTCGAAGTTCCTGGA 2091
Qy 705 PheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGln 724
Db 2092 TTTGATGATGAGAAATCACTACTGATGTCTCAATGAGCTTAGAGAAAGATTTGGTCAA 2151
Qy 725 SerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValProGlnSerSerThr 744
Db 2152 TCTAGTGTTCCTAGCCTCCACTCTGATGGAATATGGTGGTTCCTCAATCCGCAACT 2211
Qy 745 ProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSer 764
Db 2212 CCAGAATCTCTCTGAAGAGGCCATACATGTATCAGCTGTGGCTATGAAGACAAAGT 2271
Qy 765 GluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGluAspIleLeuThrGly 784
Db 2272 GACTGGGAACCTGAGATTGGATGGATCTATGTTCTGTTCAGAGATATTTCTACTGGA 2331
Qy 785 PheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetProLysArgProAlaPhe 804
Db 2332 TTCAAGATGTCATGCACGTGGCTGGCGTCAATCTACTGCATCCCTAAGCGACGATTC 2391
Qy 805 LysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeu 824
Db 2392 AAGGGGTCTGCCCAATTAATCTTCCGATCGCTGAAACAGGTGCTTCGGGGGCTCTT 2451
Qy 825 GlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyrGlyTyrGlyArg 844
Db 2452 GGTCTGTGGAAATTCCTTTTAGCAGGCATTCCTCCCAATGATGATGGGTACGAGGACGG 2511
Qy 845 LeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeu 864
Db 2512 CTTAAGTTCCTGGAGAGATTTGCTTACATCAACACCACTATTATTCACATCGATT 2571

```
QY      865  ProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMet 884
Db      2572  CCACCTCCTCTTACTGCATATATGCGAGCTATCTGCTTCTCCTCAGTGGGAAGTTTCATCAVA 2631
QY      885  ProGluIleSerAsnLeuAlaSerIleTTPheIleAlaLeuPheLeuSerIlePheAla 904
Db      2632  CCAGAGATTAGCAACTTGGAGTATTTGGTTTATCTCTCTTCTTCTGTCATCTTTGCT 2691
QY      905  ThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrpTrpArgAsnGlu 924
Db      2692  ACTGTTATCTCTTGGATGAGTGGAGTGGTGTGGCATCATCAATGATGGTGGAGGAATGAG 2751
QY      925  GlnPheTrpValIleGlyIleSerAlaHsLeuPheAlaValPheGlnGlyLeuLeu 944
Db      2752  CAGTTTGGTTATTGGAGGTATTTCCGCTCATCTATTATTTGCCGTCTTCAGGGTCTGCTG 2811
QY      945  LysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGlu 964
Db      2812  AAGGTGCTTGTGGTAATTGATACCAAGCTTCACCTGTCACCTCAAGGCCCTCTGATGAAGA 2871
QY      965  GlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuLeuProProThrThr 984
Db      2872  GCGGATTTGCTGAGCTTACATGTTTCAAGTGGCAACCCCTTCTGATCCACCGACCACT 2931
QY      985  IleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGly 1004
Db      2932  ATTTTGAATCATCAACTGGTTCGGTGTGTTGTTGCTGGTATCTCTACGCTATCAACAGTGGC 2991
QY      1005  TyrGlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrpValIleValHis 1024
Db      2992  TACCAGTCATGGGTCGGCTCTTTGGGAAGCTCTTCTTTGCCCTCTGGGTGATTGTCAC 3051
QY      1025  LeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrIleValIle 1044
Db      3052  CTGTACCCCTTCTTAAGGGTCTCATGGTTCGGCAGAACCGTACACCGACCACTTGTGT 3111
QY      1045  ValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArgValAspProPhe 1064
Db      3112  GTTTGGGCATCTCTCTGTCATGATCTTCTCTGCTGGGTGCTGATGATCATTCATTC 3171
QY      1065  ThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db      3172  ACCACCGGTGCACAGCGCCCTCATACCCCAAAATGTGGTATCACTACATGTC 3219

RESULT 8
AK069196
LOCUS
DEFINITION
ORyza sativa (japonica cultivar-group) cDNA clone:J023003G18, full
insert sequence.
ACCESSION
AK069196
VERSION
AK069196.1 GI:32979220
KEYWORDS
FLI_CDNA; CAP trapper.
SOURCE
ORyza sativa (japonica cultivar-group)
ORyza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Zhrhartoideae; Oryzeae; Oryza.
REFERENCE
1
The Rice Full-Length cDNA Consortium, National Institute of
Agricultural Sciences Rice Full-Length cDNA Project Team:
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohneda, E., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Konda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akanishi, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1. 4282
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="rRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023003G18"

ORIGIN
```

```
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 4282)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, K., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akanishi, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agricultural Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Konda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akanishi, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1. 4282
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="rRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023003G18"
```

Alignment Scores:

Pred. No.:	0	Length:	4282
Score:	5165.50	Matches:	949
Percent Similarity:	94.52%	Conservative:	68
Best Local Similarity:	88.20%	Mismatches:	54
Query Match:	89.40%	Indels:	5
DB:	8	Gaps:	4

US-09-900-237A-30 (1-1080) x AK069196 (1-4282)

QY	6	AspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIleCysAlaAspGly	25
Db	278	GACGGTGCAGAAATCGGGGAAGCAA-----TGCCATGTCTGCCAGATCTCGCGCGACGCG	331
QY	26	LeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCysArgPheProVal	45
Db	332	GTGGCGACGGCGCGGAGCGGAGCTCTTCACGCGCTCGCAGCTCTCGGGGTTCGCGGTG	391
QY	46	CysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAlaCysLeuGlnCysLys	65
Db	392	TGCGGGCATTGCTACGAGTACGAGCGCAGCAGCGCAGCGCGCTGCCCGCAGTGCAG	451
QY	66	ThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGluGlyAspAspThr	85
Db	452	ACCAAGTACAAAGCGCCCAAGGGGAGCGCCGCGATACTTGGGGATGAAAGCGATGATGT	511
QY	86	AspAlaAspGlySerAspPheAsnTyrProIleSerGlyThrGluAspGlnLysGln	105
Db	512	GATCGCGATGATGCTAGTGTGTAATCTCCACATCCGCGCAACCGAGGACCAATAGCAC	571
QY	106	LysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGlySerGlyAsnValGly	125
Db	572	AGATTGCCAGAGGATGCTCACCTGGCGCATGNACTCTGGGAGGAATGATATTTGTC	631
QY	126	HisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAspSerGlyGluIlePro	145
Db	632	CATTCTAAGTATGATAGTGGTGAATTTGGTCATCCCAAGTATGACAGTGGTGAATCCCT	691
QY	146	ArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGluIleProGlyAlaSer	165
Db	692	CGCATATATATCCCATCGCTCACTCACGACGAGATCTCAGGTGAAATTCCTGGAGCGTCC	751
QY	166	ProAspHisHisMetSerProThrGlyAsnIleSerArgArgAla---ProPhePro	184
Db	752	CCTGAT---CATATGATGTCTCCGTGGGNACATTCGCAGACGTGGGCATCCATTTCCC	808
QY	185	TyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIleGlyAsnValAla	204
Db	809	TATGTGAACCATTCACCAAAACCCATCAAGGGAGTTCCTGGTAGCCTTGGCAATGTTGCA	868
QY	205	TrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyAlaIleProMetThr	224
Db	869	TGGAAAGAGAGAGTAGAGTGGCTGGAAATGAAG---GATGAAGTGCATTCCTCATGGCT	925
QY	225	AsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThr	244
Db	926	AATGGTACTAGCATTTGCTCTTCAGAGGTCTGGGGTGGTGATGATGATGATCTACT	985
QY	245	GluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLys	264
Db	986	GACTATAACATGGAGAGCGCTTACTGAATGATGAAACACGCCAACCTCTATCGAGAAA	1045
QY	265	ValProIleAlaSerSerLysIleAsnProTyrArgMetValIleValLeuArgLeuVal	284
Db	1046	GTGCCAATTTGCTATCCAGAAATAATCCGTACAGATGGTTATGTGCTCCGATGATTT	1105
QY	285	ValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeu	304
Db	1106	GTCCTATGATATCTTGCATACCGTATCAAAATCTGTGCGTAATGCATACCCGCTA	1165
QY	305	TrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPhe	324
Db	1166	TGGTTGCTCTCTGTCATATGTGAGATTGGTTTGTCTCTATCTCTGATTCGATTCAGTTC	1225

QY	325	ProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAsp	344
Db	1226	CCCAAGTGGTCCCAATCAACCGTGAACATTTACTGGTAGGTGGCTTTAAGGTATGAC	1285
QY	345	ArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspPro	364
Db	1286	CGGGAAGGTGAACCAATCTCAATGGCTCCTGTTGACATTTTGTCACTACTGTGATCCT	1345
QY	365	LeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyr	384
Db	1346	ATGAAAGAACCTCTCTCTGTCTCACTGCCAATACTGTCTCTCCATCTCTGTGTGGATAC	1405
QY	385	ProValAspLysValSerCysTyrValSerAspAspGlyValSerMetLeuThrPheAsp	404
Db	1406	CCTGTTGATGAAGGTATCTTCTATGTCTGTATGATGAGCTGCAATGTTAACTTTGAT	1465
QY	405	AlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAsp	424
Db	1466	GCCTTGTCTGAGACTTCAGAAATTTGCTAGAAAATGGGTACCGTTCTGTAAAAAGTACAGC	1525
QY	425	IleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLys	444
Db	1526	ATGAAACCCAGAGCTCCAGAGTGTACTTTTGCACAGAAAATTTGATTACTTTGAAAGACAAA	1585
QY	445	ValGlnProSerPheValLysAspArgAlaMetLysArgGluTyrGluGluPheLys	464
Db	1586	GTTTCAGGCTCTTTTGTAAAGATCTGCTGTCATGAGAGGGAATATGAGAAATTTAAA	1645
QY	465	IleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGluGlyTrpIleMet	484
Db	1646	GTTCGTGTTAATGCTCTCGTCGCAAGGCACAGAAAGTTCTCTGAGGAAGGATGATATG	1705
QY	485	GlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnVal	504
Db	1706	CAAGATGGCACACCTTGGCTGTGTAACAATACAGAGACCATCTCTGGAAGATTTCAGGTT	1765
QY	505	PheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyr	524
Db	1766	TTCTTGGTCTATAGTGGAGCCCTTGATCCGAGGCAATGAGCTTCTCGTTAGTCTAT	1825
QY	525	ValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAla	544
Db	1826	GTGTCTCTGAGAAAACGTCTCGAATCCAAACACCAATAAAAGGCTGCTGATGAATGCA	1885
QY	545	LeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAsp	564
Db	1886	CTTGTCTGTATCAGCTGTCTTACTTAATGGCAATACTTGTGTAATCTTGACTGTGAT	1945
QY	565	HisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsn	584
Db	1946	CATCATCAACAATAGCAAGCTCTCCGAGAGGCTATGTGCTTCTTATGGATCCAAAC	2005
QY	585	LeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsn	604
Db	2006	CTAGGAGGCTGTCTGTTATGTCCAATTCCTCAGAGGTTGACGGTATCGATAGAAAT	2065
QY	605	AspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGly	624
Db	2066	GATCGATATGCAACACAGAAACACCGTGTTCATATTAATTTGAGAGGTTCTCGATGGT	2125
QY	625	IleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGly	644
Db	2126	CTCCAAGACCAAGTTATGTGGGAACCGTGTGTGTTCACAGACAGCTCTTTATGGT	2185
QY	645	TyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeuCysGlyLys	664
Db	2186	TATGAACCCCCCAATTAAGCAGAGAGGCTTACTTCTTCCTCGCTTTTGGGGGACGA	2245
QY	665	LysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisVal	684
Db	2246	AAGNAGCAAAAAGTCCGAGGAGAGACGCAAGAAAAGAAAAGTACACAAACATGTG	2305

Percent Similarity: 93.99%
 Best Local Similarity: 87.60%
 Query Match: 87.01%
 DB: 8

Conservative: 69
 Mismatches: 59
 Indels: 6
 Gaps: 6

US-09-900-237A-30 (1-1080) x AF200529 (1-3676)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
 DB 254 ATGACGGC--GCGCAGCGCCACGAAATCGGGGAAGCATGTGGCGGCGAGGTGTGCCAG 310
 QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
 DB 311 ATCTGCGGCACGGCGTGGACCCGCGCGGAGCGCGACCTTTCACGGCTCGCAGCGTC 370
 QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
 DB 371 TGGGGCTTCGCCGTGCGGCCCATGCTACAGTACAGCGCAGGACGCCACCCAGCGC 430
 QY 61 CysLeuGlnCysLysThrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
 DB 431 TGCCCCAGTGCAGAGTAAGTACAAGCGCCACAAGGGAGGCCACCCAGTACACGGGTGAG 490
 QY 81 GluGlyAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
 DB 491 GAAATGAGATGTGGATGCTGAGATGTGAGTACTACAACCTACCAAGCATCTGGCAAC 550
 QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
 DB 551 CAGGATCAGAAAGAAAGATTCTGAGAGATGCTCACTTGGCGGCAAACTCACGTGGC 610
 QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
 DB 611 AGT---GATATTGGCTGGCTAAGTATGACAGCGGTGAAATTTGGGCATCGGAAGTATGAC 667
 QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
 DB 668 AGTGTGAGATCCCTCGTGATATATCCGTCACCTAACTCATAGCCAGATCTCAGGAG 727
 QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
 DB 728 ATTCCTGGAGCTTCCCTCGAT---CATATGATGTCTCTCTGTTGGGAAATTTGGCAGGCGT 784
 QY 181 Ala---ProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySer 199
 DB 785 GGACATCAATTCCTTATGTAAATCATTTCTCAACCCATCGAGGAGTCTCCGGTAGC 844
 QY 200 IleGlyAsnValAlaTyrLysGluArgValAspGlyTyrLysMetLysGlnAspLysGly 219
 DB 845 CTTGGCAATGTTGCATGGAAGAGAGGGTGGATGGATGGAAATGAAG---GATAAAGGT 901
 QY 220 AlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAsp 239
 DB 902 GCAATTCCTATGACCAATGGAACAAGCATTCCTCCATCAGAAGGCGGTGGAGTTCGTGAT 961
 QY 240 IleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGln 259
 DB 962 ATTGATGCTTCTACTGATTAATTAACATGGAAGATCCCTTACTGAAATGATGAATCGGCAA 1021
 QY 260 ProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIle 279
 DB 1022 CCTCTATCTAGAAAAGTGCCAAATTCCTTCATCCAGAATAATCCGTACAGAAATGGTCATT 1081
 QY 280 ValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArg 299
 DB 1082 GTCTACGTTTGGCTGTCTATGCATATCTTGGCTACCGTATCCATCATCTGTGAC 1141
 QY 300 AsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrp 319
 DB 1142 AATGCATATCCACTGTGGCTTTTATCCGTATATGAGATCTGGTTGTGTTGCTCTG 1201
 QY 320 IleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeu 339

DB 1202 ATTTTGGATCAGTTCCCAAGTGTGTCCTCCCAATCAACCGTGAAACATACCTTTGATAGACTG 1261
 QY 340 AlaLeuArgTyrAspArgGlyGluProSerGlnLeuAlaAlaValAspIlePheVal 359
 DB 1262 GCTTTAAGGTATGACCGAGAGGTGAACCATCTCAATTAGTCTCTGTTGATATTTTGTGTC 1321
 QY 360 SerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIle 379
 DB 1322 AGTACTGTGATCAATCAATAGGAGGCTCTCTGTCTACTGCAATATCTGTGCTTCCATC 1381
 QY 380 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer 399
 DB 1382 CTTGCTGTCAATATCCGTTGACAGGTATCTTGTCTATGTTTGGATGATGGAGCTGCT 1441
 QY 400 MetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPhe 419
 DB 1442 ATGCTGACTTTTGTATGCTCTCTGAAACTTCAGAGTTTGTAGAAAATGGTTCCTGTC 1501
 QY 420 ValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAsp 439
 DB 1502 TGTAAGAAGTACAAACATAGAGCCCTANGGCCCCGGAATGGTACTTTGCTCAGAAAATTGAT 1561
 QY 440 TyrLeuLysAspLysValGlnProSerPheValLysAspArgArgAlaMetLysArgGlu 459
 DB 1562 TACTTGAAGACAAAGTTCAAACCTCATTTGTGAAGAAGCGCGGCCCATGAAGAGAGAA 1621
 QY 460 TyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGlu 479
 DB 1622 TATGAAGAATTCAAAGTTCTGTATCAATGCTCTGTAGCCCAAGCCACAAAAGTTTCCGAG 1681
 QY 480 GluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisPro 499
 DB 1682 GAGGATGATCATGCAAGATGGTACACCTTGGCTGGGAAACAATACTAGGACCATCTCT 1741
 QY 500 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeu 519
 DB 1742 GGAATGANTCAGTWTTCCTGGGTCAAGTGGAGGCTTGACGTTGAAGGCAATGAACAT 1801
 QY 520 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAla 539
 DB 1802 CCTCGTTTGGTTTATGTGTCTCGTGAATAAGCTCTCGATTCACACATCACAAGAGGCT 1861
 QY 540 GlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeu 559
 DB 1862 GGTGCCATGAATGACCTTGTGTGTATCAGCTGTCTTACTAATGGGCAATACATGTTG 1921
 QY 560 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPhe 579
 DB 1922 AATCTTGATTGTGACCACTACATCAATAATAGCAAGGCTCTTCGAGAAGCTATGTGCTTC 1981
 QY 580 LeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAsp 599
 DB 1982 CTTATGACCCCAACCTTAGGAAGGAATGTCTGTATGTCCAATTTCTCAGAGGTTTGTAT 2041
 QY 600 GlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeu 619
 DB 2042 GGTATTCATAGAAATGACCATATGCAAAACAGGAACACTGTGTTTTTCGATATTAACAT 2101
 QY 620 ArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArg 639
 DB 2102 AGAGGCTTCGACGGCATTCAGGGCCAGTTTATGTGGGAACCTGTTGTGTGTTTAAACAGA 2161
 QY 640 ThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSer 659
 DB 2162 ACGGCTTATATGTTATGAGCTCCAGTCAAGAAAAAAGCAGGCTCTCTCTCTCTCT 2221
 QY 660 LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLys 679
 DB 2222 CTTTGTGGGGAGGAAAAGAGCTCAAAATCTAAGAAG---AGCTCGGAAAAAGAAAG 2278
 QY 680 SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGluGly 699
 DB 2279 TCACATAGACGCGACAGTCTCTGTACCAGTATTATTAATCTCGAAGATATAGAGGAGG 2338

700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2339 ATTGAAGTTCTCAGTTTGTGATGATGAGAAATCGTGAATTATGTCTCAATGAGCTGGAG 2398
720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyVal 739
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2399 AAGAGATTGGCCAGTCCAGTGTGTTGTAGCCTCTACTCTGATGGAAATATGGTGGTGT 2458
740 ProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGly 759
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2459 CCAAAATCTCCAACTCCAGAGTCTCTCTGAAAGAAGCTATTCATGTCAATCAGCTGGC 2518
760 TyrGluAspLysSerGluThrGlyThrGluIleGlyTrpLeuTyrGlySerValThrGlu 779
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2519 TATGAGGACAAACTGACTGGGAATGAGATTGGTGGATCTATGCTTCTGTTACAGAA 2578
780 AsnLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetPro 799
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2579 GACATCTCACCGGATTCAAGATCATGCTCGAGCTGGCGATCAATCTACTGCTATGCT 2638
800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2639 AAGCGACAGCTTTCAAGGATCTGCTCCATCAACCTTCGGATCGTTTGAATCAAGTG 2698
820 LeuArgTrpAlaLeuGlySerValGluLeuPheSerArgHisCysProLeuTrpTyr 839
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2699 CTTCGGTGGCTCTTGGTTCCATTGAAATCTTTTTCAGCAGGCATTTGCCATGATG 2758
840 GlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2759 GCCTATGAGGCGCGCTTAAATTTCTCGAGAGATTGCTTATATCAACACAAATTTAT 2818
860 ProLeuThrSerLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2819 CCATCTCATCAATCCCCGCTCTCTGTAAGTATGATGATGATGATGATGATGATGAT 2878
880 GlyLysPheIleMetProGluLysSerAsnLeuAlaSerIleTrpPheIleAlaLeuPhe 899
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2879 GGGAAATTCATCATCCCAAGATTAGTAACCTAGAGAGTGTGTTGTTTATATCGCTCT 2938
900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGlu 919
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2939 ATCTCAATCTTGGCCACTGGTATCTCTGAGATGAGGTGGAGTGGTGGATGATGAA 2998
920 TrpTrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaVal 939
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
2999 TGGTGGAGAACGAGCAGTCTCGGTCATGTTGGTATTTCTCGGCATTTATTTGCCGTC 3058
940 PheGlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 959
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3059 TTCCAGGGTCTCTGAAGTGTGTTGTTGATTCGACACGAGCTTCACTGACCTCTAAG 3118
960 AlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeu 979
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3119 GCCACTGACGAAAGAGTGAATTTGCGGAGCTCTCATGTTCAAGTGGACACACCTTCG 3178
980 IleProThrThrThrIleLeuIleLeuAsnMetValGlyValAlaGlyThrSerTyr 999
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3179 ATCCCAACCAACCACTATTTTATCATCAACCTGCTGGCGTGGTGGCTGTCATTTCT 3238
1000 AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPhe 1019
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3239 GCATCAATAGCGTTTACCACTCATGGGAGCTCTTTTCGGGAGGCTCTCTTTTGGCTTC 3298
1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3299 TGGGTGATTTGCCACTGTACCTTCTCAAGGCGCTCATGGGAGGAGGAGGAGGAGGAG 3358
1040 ProThrIleValIleValTrpAlaValLeuAlaSerIlePheSerLeuLeuTrpVal 1059
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3359 CCGACCAATGTCGTGTGTCGGCTATCTCTTGGCTGATCTTTTCCCTGATGTCGGTT 3418

FEATURES
Location/Qualifiers
source

1. .87792

1060 ArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3419 CGTATCGATCCATTCACCCCGGTCCTGCGCCTGATATCGGAAATGTGGCATCAAC 3478
1080 Cys 1080
|||
3479 TGC 3481
RESULT 10
AP003837/c
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,
BAC clone: OJ1559_F09.
ACCESSION
AP003837
VERSION
AP003837.2
KEYWORDS
GI:26017216
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
clone: OJ1559_F09
Published Only in Database (2001)
2 (bases 1 to 87792)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (03-JUL-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kamondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasakienias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
On Dec 3, 2002 this sequence version replaced gi:14595177.
Genes were predicted from the integrated results of the following:
GENSCAN (<http://ccr-081.mit.edu/GENSCAN.html>), FGENESH
(<http://www.softberry.com/>), GeneMark.hmm
(<http://opal.biology.gatech.edu/GeneMark/>), GlimmerM
(http://opal.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM
(<http://rgp.dna.affrc.go.jp/RiceHMM/>), SplicePredictor
(<http://bioinformatics.iastate.edu/cgi-bin/sp.cgi>), Sim4
(<http://www.tigr.org/software/glimmer/>), ELASTN and BLASTX. The
genomic sequence was searched against NCBI Nonredundant protein
database, nr (<ftp://ncbi.nlm.nih.gov/blast/db>) and the cDNA
sequence database at RGP or DDBJ. Protein homologies of the coding
regions were searched against NCBI Nonredundant Protein database
with BLASTP. ESTs represent the identified cDNA sequences using
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DDBJ accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to RGPSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.
The orientation of the sequence is from M13rev to -21M13 of the BAC
clone. This sequence of OJ1559_F09 clone has an overlap with
OJ1119_B04 (DDBJ: AP003943) clone at 5' end and with OJ1136_A05
(DDBJ: AP003748) at 3' end. The sequence was generated by combining
Monsanto and RGP-Japan sequencing data. Detailed information on
overlap and assembly quality together with annotation of this entry
is available at
<http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

```

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="7"
/clone="OJ1559_F09"
complement(2210..2833)
/gene="OJ1559_F09.101"
complement(join(<2210..2425,2690..>2833))
/gene="OJ1559_F09.101"
/notes="start and end point are not identified"
complement(join(2210..2425,2690..2833))
/gene="OJ1559_F09.101"
/notes="predicted by FGENSEH etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAC57272.1"
/db_xref="GI:28411797"
/translation="MAADAPMLGTMVTVEETPYDLDPPTSKLDADHQGFTTTRPCR
TKDVLFSITFAWVTVYVCLVHDDLSLQTLFELVWVVIQWLIVFVMTYSAPDDAN
QEYRKRYNVVNAWFC"
complement(join(4610..4836,5286..5394))
/gene="OJ1559_F09.102"
complement(join(4610..4836,5286..5394))
/gene="OJ1559_F09.102"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(join(6618..6632,7157..7330))
/gene="OJ1559_F09.103"
complement(join(6618..6632,7157..7330))
/gene="OJ1559_F09.103"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(7690..9417)
/gene="OJ1559_F09.104"
complement(join(<7690..7908,7998..8513,8614..>8634,
8736..8906,8992..9123,9208..>9417))
/gene="OJ1559_F09.104"
/notes="start and end point are not identified"
complement(join(7690..7908,7998..8513,8614..8634,
8736..8906,8992..9123,9208..9417))
/gene="OJ1559_F09.104"
/codon_start=1
/product="putative polygalacturonase Pg1"
/protein_id="BAC57273.1"
/db_xref="GI:28411798"
/translation="WACNVLIISAAYVVMVMSFLLCSVHAKVNNASPHLTSGDDSLRGR
SLESKVDFVRKHGAYGGQDDTKALAKAAWAAACSSQPSIVLIPKGRYLTKHIT
LSGPCSSITFMIEGTLVAPKRSWMSKETIRHWMFNGVSLTVAGGCTVDENGKIW
WQNSCKTNAKLPCTAPALTALFYSCNLKLVNLLKNSQIHMVSVEDCTDVRISLTI
TAPGTPNTDGIHITRSKNVQVTCIITKGDGMSIEDGTENLHVKNVCGPGHGISI
GSLGDHSEAHVNVTVTRLYGTNGARKITWQGGNGYAKNIYVQNMWENWNPPI
LIDQNVCDSPATCKEQTSAVQVSNVVFNRKIRGTSAKKAIKLDCSRNVPCQGITLKV
KLITKGGSDAKSTCGNAKWKSGIALCFQ"
join(10069..10162,10373..10608,11858..12025)
/gene="OJ1559_F09.105"
join(10069..10162,10373..10608,11858..12025)
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
12214..17342
/gene="OJ1559_F09.106"
join(12214..12339,12451..12516,13288..13553,14328..14452,
14845..14938,15622..16079,16691..17342)
/gene="OJ1559_F09.106"
/notes="supported by full-length cDNA(s): AK109769"
join(13400..13553,14328..14452,14845..14938,15622..16079,
16691..16966)
/gene="OJ1559_F09.106"
/notes="contains full-length cDNA(s): AK109769"

/codon_start=1
/product="NLI interacting factor-like protein"
/protein_id="BAC83123.2"
/db_xref="GI:38175433"
/translation="MOTRKKGATKNALGDPASTTKTSRQPRAAQAAASEKKVNDLITS
SAKKKSGVAPSKKNRASKGRKLIISACDAANSENVSVQVGGIPHOQSDNDVGRP
CNSIFSPAYHLQKECGASNFAGKLEHGRDITLGSVSSVEERTTHAQGRKVTTSRST
SHAVKTCVGDHHTLNAQSAFCNTPLFEDEFSELGNLSSEVSAIYLAMQSKLECIDF
HSQDSITEGVDPDEEDDFDYAFIKDLPDLSLVKPKFRPVLIPKQTRSCPITTT
LKHLICILNARMILTSHFQIFILTSNTQVMSDAVPIKSKSWRGLPICLRQSFSLQVK
AFMQNSFMFLILKESCSVIVFTVIVPMRETT"
complement(19106..20743)
/gene="OJ1559_F09.107"
complement(join(<19106..19339,19456..19971,20079..>20270,
20360..20491,20600..>20743))
/gene="OJ1559_F09.107"
/notes="start and end point are not identified"
complement(join(19106..19339,19456..19971,20079..20270,
20360..20491,20600..20743))
/gene="OJ1559_F09.107"
/codon_start=1
/product="putative polygalacturonase precursor"
/protein_id="BAC57275.1"
/db_xref="GI:28411800"
/translation="MASTVLIIVSASLLALFFFLHSADADVGSNVFSIQSYGAHGDGRH
DOTKALGDTWAACSSAKPAVLIPKGRKYLTKHITLSPGKSSISLMVKGSLVASPE
RDSWAKETIRHMLISGVTGLTVGGTIDNGKIWMQNSCKTNSKLPCTEAPTALTFF
YSCNLKLYLVKVNQQIQISVEDCTVMVSRLSITAPETAPNTDGIHITRDRVEV
TDCMIKTGDDCMSIEDGTENLHVKNVCGPGHGISISGLGDHSEAHVNVTVNDVRL
YGTANGARIKTWQGGKSAKNIYVQNMWMDVNPPIIDQNYCDSFTPCQKQSAVEV
SNVLPKNIRGTSASEAIMLHCSSTVPCGHITLNVNLIVKGGSSNAKSTCQNAEWKK
SGSVSPQPCGFRN"
24140..25351
/gene="OJ1559_F09.108"
<24140..>25351
/gene="OJ1559_F09.108"
/notes="start and end point are not identified"
24140..25351
/gene="OJ1559_F09.108"
/notes="similar to Oryza sativa chromosome7, P0418E08.23"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAC57276.1"
/db_xref="GI:28411801"
/translation="WAPPRKRRRVDSGGNGNYPVLVATPNDGVLVLDLNAVLRL
PASPACRLRAVCRPWRAVLSDPRFAAHAARHPDLHVVAACDLDDAGGIELVDVLY
GASGDVAKRVFAGRCDDTDAVSARDGVALLVGNDRRLVLDAAASAAVSLVPDGGHPH
INCSFTLGRAASSSGCEHKLVRIGTVVHGEPOVCAVLTAVAGRGQNARWRAPSPPL
VVRTRRGVAVAGGVAVYLLRRAYLADRIAAPDLAEQWRPALVGPPFLAAWRPFRPD
RPRVTLAELGSLVLAIDHRAATLIDMLFLAAGDEQHWKSKQYTTVTPVHRRPWRCD
GERAPVVLDDGRILVFWWAGSGGTHGGGVWVRVDPITGGHTDVAARCAHVGV
YTGNNLSLVSE"
complement(join(26590..26940,27083..27243,27365..27502,
28511..28604))
/gene="OJ1559_F09.109"
complement(join(26590..26940,27083..27243,27365..27502,
28511..28604))
/gene="OJ1559_F09.109"
/notes="supported by full-length cDNA(s): AK062943"
complement(join(26842..26940,27083..27243,27365..27497))
/gene="OJ1559_F09.109"
/notes="contains EST(s): AU089698(C2176),D23094(C2176)
contains full-length cDNA(s): AK062943"
/codon_start=1

Alignment Scores:
Pred. No.: 0 Length: 87792
Score: 4838.00 Matches: 991
Percent Similarity: 62.37% Conservative: 50
Best Local Similarity: 59.38% Mismatches: 37
Query Match: 83.73% Indels: 596
DB: 8 Gaps: 14

```

US-09-900-237A-30 (1-1080) x AP003837 (1-87792)	
QY	3 GlyAspAlaasp---AlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIle 21
Db	64877 GGTGATCTATCTGTTGGCAGAGAGTCCGGAGGACACGGAGCGGCGAGGTCACAGATC 64818
QY	22 CysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCys 41
Db	64817 TGGCGGACGGCGTGGGACACGACGGGAGGCGAGCTGTTCGCCGCTCGCAGCTGC 64758
QY	42 ArgPheProValCysArgProCysTyrGluHisGluArgGlyGlyThrGlnAlaCys 61
Db	64757 GGCTTCGCCGTGTCGCCCTCTGTCAGAGTCAGGCGAGGATGGCACCAGGCTTC 64698
QY	62 LeuGlnCysLysThrLysTyrLysArgHis-Arg- 72
Db	64697 CCCAGTGCAGACCAAGTACAGAGGCCACAGGGTAAATGCTACCATCTCTCTCTC 64638
QY	72 ----- 72
Db	64637 CTCTCCCTCTCTCCATTCCTCCCGATCTCCGCCCATCGACTGTAAATAAGTTCAGAAA 64578
QY	72 ----- 72
Db	64577 CATCTCAGAAAAAAGTCTGGATCTCTAGTTCAGAACAGTCAAAACGAGCATTCCTAT 64518
QY	72 ----- 72
Db	64517 AGATGGATCCTGTTGCTTCTTGTAGGAAGGTCATACAGATCCAATGTTTGTGATTT 64458
QY	72 ----- 72
Db	64457 TTTGGAAATTCAGAAAAAATAATGTCGATTAGTAAACGATGTTACAGTCATGCTAAGGC 64398
QY	73 -----GlySerProAlaIleArgGlyGluGlyAs 83
Db	64397 TTTGCTGAACGTGTATGGAAATTAATCAGGAGGCCCGCGATCCGTGGGAGGAAGGGA 64338
QY	83 pAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThrGluAspG 103
Db	64337 GGATCTGATGTCGATGATGTCAGTGACTACAACTACCTGCTGTCATCGCAGTGGCGACA 64278
QY	103 nLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlySerGlyAs 123
Db	64277 GAAGCAGAAGATTGCTGATAGGATCGCAGTTGGCGCATGAATGCTGGGGGTGGGAGA 64218
QY	123 nValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAspSerGlyG 143
Db	64217 CGTCGCCGCTCCCAAGTATGACAGTGGCGAGATCGGCTCACCAAGTATGACAGTGGCGA 64158
QY	143 uIleProArgGlyTyrValProSerValThrAsnSerGlnMet- 157
Db	64157 GATCCTCGGGATACATCCCTTCAGTCACTAATAGCCAGGT-GGGTGGCAGTTTTTAAC 64099
QY	157 ----- 157
Db	64098 AGATCGTGATACAGTGTTCGGGTATGCTCAAAATGAATTAATCTTGAAATGTTTGGCA 64039
QY	158 ---SerGlyGluIleProGlyValaSerProAspHisHisMetMetSerProThrGlyAs 176
Db	64038 GATCTCGGAGAAATCCCTTGGAGCTTCCCTGATCATCATGATGATGCTCTCTACCGAAA 63979
QY	176 nIleSerArgAlaProPheProThrValAsnHis----- 188
Db	63978 CATTCGCAAGCGTCTCCATTCCTTATGTGAACCATTCAGCATGTTTATTTCTTGATA 63919
QY	188 ----- 188
Db	63918 TGATGTCACCTCTCAATTAGGTACATCCGCTTGAGATCTAAGATCGTAATPATCTTTCTT 63859
QY	189 -----Ser-ProAsnProSerArgGluPheSerGlySerIleGlyAsnValAlaTrpL 206
Db	63858 GCTTCTTCTAGCAAAATCCATCAAGGGAGTCTCTGGCAGCATGGAAATGTTGCCTGGA 63799
QY	206 ysGluArgValaAspGlyTrpLysMetLysGlnAspLysGlyAlaIleProMetThrAsnG 226
Db	63798 AAGAAGAGTGTGATGCTGGAAACTGAAGCAGACAAAGGAGCAATTCCTCATGACCAACG 63739
QY	226 lyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThrGluT 246
Db	63738 GGACAAAGCATTCGCCCTTCTGAAGGTCGGGAGTTGGTGATATCGATGATCCATCGATT 63679
QY	246 yzAsnMetGluAspAlaLeuLeu----- 253
Db	63678 ACAATATGAAGATGCTTACTGTGAGTCTGCATCTCCCTCAGGAGACATAATCTT 63619
QY	253 ----- 253
Db	63618 ATACGGGATCATATCGCTGACATTTACAAATGTAGTATGTGTGATTTCTATCTATAACA 63559
QY	253 ----- 253
Db	63558 CTGACTAAACCGAAAAATTATTATATCGATGTTATTGTTTTTCAACATAGACATCAACATA 63499
QY	253 ----- 253
Db	63498 TAGCCAAAGCAAGTGGGTGACCTAAATGTATTAGAGCTGCACITTTTTTCCACTAACACTT 63439
QY	253 ----- 253
Db	63438 CCAGCTATAATTTTTGGTACATCGAATTGGTACTAAACAAAAAGAGGTGTCTGCTCTAAG 63379
QY	253 ----- 253
Db	63378 TTCCTAGTCAACTCATTAGTGTGCTGCTGCTGAGGTGAATGAGGGTTTTTGTCTCCTGAT 63319
QY	253 ----- 253
Db	63318 AGAAGGAAGCAAAATAATATCTGAAAGTTACAAATGGAATAATTAGAGATTAAATAT 63259
QY	253 ----- 253
Db	63258 GTGACATTTTTGATGGTGTCATAACAAGTCTGTGGCATTAACCTTTATGCAATTTGAATT 63199
QY	254 ---AsnAspGluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysI 272
Db	63198 GCAGGAATGATGAAACTCGCCAGCCTCTCTCTAGAAAAAGTTCCCTCTCTCCATCCAGAA 63139
QY	272 leAsnProTyrArgMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisT 292
Db	63138 TAAATCCCTACAGAAATGTCATGTTCTGCGTTTGGTTGTTCTTAAGCATTTTCTTCACT 63079
QY	292 yrArgLeuThrAsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysG 312
Db	63078 ACCGTATTACGAATCCTGTGCGCAATGCGTATCCGCTCTGGCTTTTATCTGTATATATG 63019
QY	312 luIleTrpPheAlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnA 332
Db	63018 AGATTTGGTTTCTTGTCTGCTGATATTGGATCCGGAATGAGTCCGGAAGTGGTTTCCAATCAAC 62959
QY	332 rgGluThrTyrlleuAspArgLeuAlaLeu----- 341
Db	62958 GTGAACCTTACCTTGATAGGCTGGCATTGAGGTGATCTCCTTTTAGTTCTGTGTGACTT 62899
QY	341 ----- 341
Db	62898 AATCCAAACAAACATCCCTTATTCAATTTGCTGATGCGGTGTACCATCTCATTTTCCATTAC 62839
QY	342 ArgTyrAspArgGlyGluProSerGlnLeuAlaAlaValAspIlePheValSerThr 361
Db	62838 AGGTATGACAGAGAGGTGAGCCATCTCAGTTGGGTGCTGTGTGACATTTTGTTCAGTACA 62779
QY	362 ValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeuAla 381
Db	62778 GTCGACCCCATGAAGGAACCTCTCTTGTGTACTGCCAATACCGTGTCTGCTTCTTCTTGT 62719

```
QY 382 ValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeu 401
Db 62718 GTTGATTACCCAGTGGCAAGAGTCTCTTGCTATGATCTGACGATGGTGGCTGCAATGCTG 62659
QY 402 ThrPheAspAlaLeuAlaGluThrSerCluPheAlaArgLysTrpValPropheValLys 421
Db 62658 ACTTTTGATGATTTGGCTGAGACTTCAGAGFTTGCTAGAAAGTGGGTTCATTCGTTAAG 62599
QY 422 LysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeu 441
Db 62598 AAGTATAACATTGAGCCAGAGCTCTGAGTGGTACTTCTCCAGAAAATCGATTACTTG 62539
QY 442 LysAspLysValGlnProSerPheValLysAspArgAlaMetLys----- 457
Db 62538 AAAGACAAAGTCCACCCTTCATTGTGTTAAAGACCGTGGTCCATGAA-GGTTTGTGTTGCT 62480
QY 457 ----- 457
Db 62479 TTATAACACATTAAATTAATTTTGCTATCTCTGCTCCAAATGCTAACAGTCTTCCCAA 62420
QY 458 -----ArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuVal 471
Db 62419 AAAAAAAAACTGTTACAGAGAGATATGAAGAATTCAAAGTTAGGATAAATGCCCTTGTT 62360
QY 472 SerLysAlaLeuLysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrpPro 491
Db 62359 GCTAAGGCCAGAAAATGCCGAGGAGGATGATCATGCAAGATGCGACACCATGGCCA 62300
QY 492 GlyAsnAsnThrArgAspHisProGlyMetIleGlnVal----- 504
Db 62299 SGAAACAATACTAGGACCATCTCGAATGATTCAGGT-TAGTGTATTTCAAATAAAATC 62241
QY 504 ----- 504
Db 62240 TACTATGACTATGTTGATGATAGTCTTTTACTTGATGATATGATTATGTTTGCTG 62181
QY 505 -----PheLeuGlyHisSerGlyGlyLeuAspThrGlu 515
Db 62180 ATTGCTCTTTCTGATTTTCTTAGGTTTTCTTGTTGTCACAGTGTGGCTTGATACCGAG 62121
QY 516 GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis 535
Db 62120 GGTAAATGAGCTTCCCGCGCTGCTTACGTTATCTGTGAGAACGCTCGTGGGTTTCAGCAC 62061
QY 536 HisLysLysAlaGlyAlaMetAsnAlaLeu----- 545
Db 62060 CACAAAGAGGCTGTGTCATGAATGCTCT-TGTGAGTAGTTGAATATTTACAAAACCTTAC 62002
QY 545 ----- 545
Db 62001 ACTATCTAGTTCAATTTTCATTTCTCTCTTAAAGTTCTAATGAGGATTAAATCTTTTCTG 61942
QY 546 ---ValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAsp 564
Db 61941 CAGGTTCTGTGCTCAGCTGTGCTTACCAATGGACAGTACATGTTGAAATCTTGATTTGAT 61882
QY 565 HistyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsn 584
Db 61881 CACTACATCAACAACAGCAGGCTCTGGGGAAGCTATGTGCTTCTTATGGATCCAAAC 61822
QY 585 LeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsn 604
Db 61821 CTAGGAAGGAGTGTCTGTATGTTTCAGTTCCCAAAAGGTTCSGATGGTPAATTGATAGGAAT 61762
QY 605 AspArgTyrAlaAsnArgAsnThrValPhePheAspIle----- 617
Db 61761 GATCGATATGGAACAGGAACACTGTGTTTTCGATGT-AAGTTGGCAATTCCTCCTTTAA 61703
QY 617 ----- 617
Db 61702 GATTCTCGCTCATCTGCAACTACTAGTATAGTGTAAATCATATTGATTTAAACAAATGATG 61643
QY 617 ----- 617
```

```
Db 61642 CTAATATATTTTGTGATGCACTAAATTTTAAATTTCTAACAAATTCATATATTTTTTTGGGCA 61583
QY 618 -----AsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyr 630
Db 61582 TATTTCTGTCTTCTACAGATTAACTTGGGGGTCTTGATGCATCCCAAGGCCAGTTTAT 61523
QY 631 ValGlyThrGlyCysValPheAsnThrAlaIleTyrGlyTyrGluProProIleLys 650
Db 61522 GTGGAACTGGTGTGTATTCAACAGAACTGCTCTATATGTTATGAACCCCAATTAAG 61463
QY 651 AlalysLysProGly---PheLeuAlaSerLeuCysGlyGlyLysLysAlaSerLys 669
Db 61462 CAGAGAGAGAGGAGAGTTTCTTGTCTACTATGTGGGGCGAGAAAGAGCAAGCAAG 61403
QY 670 SerLysLysArgSerSerAspLysLysSerAsnLysHisValAspSerSerValPro 689
Db 61402 TCAAGAGAGAGAGAGTTCCGGACAGAGAACTCAACACAGCAGCTGGACAGTGTGTGCCA 61343
QY 690 ValPheAsnLeuGluAspIleGluGlyValGlu----- 701
Db 61342 GTTTTCAATCTTGAAGATATAGAGAGGGTGTTCGAAGGTACCATTCATATGTTCCAGTTA 61283
QY 701 ----- 701
Db 61282 TTGCATAATTTAGGANTATAAGTTGTAATCAACATATATGTCATCTCGACATATCATCTAT 61223
QY 702 -----GlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db 61222 ATGGCAGGTGCTGGATTGATGATGAGAAATCACTTCTTATGCTCTCAATGAGCTTGGAG 61163
QY 720 LysArgPheGlyGlnSerAlaIalaPheValAlaSerThrLeuMetGluTyrGlyVal 739
Db 61162 AAGAGATTTGGCCAGTCTCGACGCTTGTGTGCTCCTCATGTAATATGTTGCTGTT 61103
QY 740 ProGlnSerSerThrProGlnSerLeuLeuLysGluAlaIleHisValIleSerCysGly 759
Db 61102 CCTCAATCTGCAACCCAGAACTCTTTTGAAGAAGAGCTATCCATGTGATAAGTTGTGTC 61043
QY 760 TyrGluAspLysSerGluTrpGlyThrGlu----- 769
Db 61042 TATGAGGACAGACCCGAAATGGGGAACATGAGGTAACATGGGCTTCTCACAGCAACATTGG 60983
QY 769 ----- 769
Db 60982 TTACATCTTTTTTTAGTAACTTGTGTAATTTTACACCGGCTATGCACCTTCAATGGTA 60923
QY 770 -----IleGlyTyrIleTyrGlySerValThrGluAspIle 781
Db 60922 TGTTAATGCTGACATATGTTGCAGATTGGTGGATCTACGGTTCCTGACAGAAATATT 60863
QY 782 LeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetProLysArg 801
Db 60862 CTCACGTGATTCAAGATGCATGCGCGAGCTGGAGATCAATCTACTGATGCCAAGCGC 60803
QY 802 ProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArg 821
Db 60802 CCAGCTTTCAGGGGTCTGCTCTTATCAATCTTTTCAGATCGTCTTTAAACCAAGTGTCTCG 60743
QY 822 TrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGlyTyr 841
Db 60742 TGGGCACCTTGTTCTGTTGAAATCTTTTCAGTCCCATTTGCTCCATATGATGTCGGCTAT 60683
QY 842 GlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLeu 861
Db 60682 GGAGAGCGCTTAAAGTTCTTTGGAGAGATTTGCCTACATCAACACCATATTATTCATTG 60623
QY 862 ThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGlyLys 881
Db 60622 ACATCGATCCGGCTTCTCATATACTGTGTTTGTGCTCTATCTGCTTTGCTCCTGGAAG 60563
QY 882 PheIleMetProGluIleSerAsn----- 889
Db 882 PheIleMetProGluIleSerAsn----- 889
```

Db 60562 TTTCATATCCAGAGTAAGCAATGAGTCTGTTATACATACCTTCTATAGAAAATGTTTG 60503
QY 890 -----Leu-AlaSerileTr 894
Db 60502 AGACAACCAACTAATAATTTTGAATGCAGATTAGCAACTTTGCTAGTATTGG 60443
QY 894 pHeilleAlaLeuPheLeuSerilePheAlaThrGlyLeuLeuGluMetArgTrpSerGl 914
Db 60442 GTTCACTCTCTCTTCATTTCAATTTTTCACCTGTTATCTCTTGAGATGAGGTGGAGTGG 60383
QY 914 yValGlyIleAspGluTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAl 934
Db 60382 TGTGTGCATCGATGAGGTGGAGGAATGAACAGTTCTGGTTATTGGAGTATATCTGC 60323
QY 934 aHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPh 954
Db 60322 GCATCTTTTGGCGTCTTCAGGGTCTCTCAAGGTCTTGTGTTATGCACACCAATTT 60263
QY 954 eThrValThrSerIysAlaAsnAspGluGluGlyAspPheAlaGluLeuTrpMetPheLy 974
Db 60262 CACTGTCACTCAAGGCTTCTGATGAAGATGGCGACTTTGCTGAGCTCTACATGTTCAA 60203
QY 974 sTrpThrThrLeuLeuLeuProProThrThrIleLeuIleIleAsnMetValGlyValVa 994
Db 60202 GTGGACAACGCTTCTCATCCACCGACGACCACTTGTATCATTAACCTGCTGGTGTGTGT 60143
QY 994 lAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLy 1014
Db 60142 TGCTGGTATCTCATACGGATCAACAGCGGTACCAATCATCTGGGACCGCTCTTTGGCAA 60083
QY 1014 sLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGl 1034
Db 60082 GCTCTTCTTCCCTCTGGGTGATTCACCTTGTACCCCTTCTCAAGGGTCTTATGGG 60023
QY 1034 yArgGluAsnArgThrProThrIleValIleValTrpAlaValLeuAlaSerIlePh 1054
Db 60022 TCGGCAAAACCGCACTCCGACCATCTGTTGTTGTTGGCAATCCTCTGGCTTCACTT 59963
QY 1054 eSerLeuLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGl 1074
Db 59962 CTCATTGCTGGTTCGATCGATCCATTACACCGGTGTCCACCGGCGCCAGATACCCA 59903
QY 1074 nThrCysGlyIleAsnCys 1080
Db 59902 AACATGTGGCATCACTGC 59884
RESULT 11
AC145384/c
LOCUS
DEFINITION
AC145384
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 41693)
Buell,R.
Direct Submission
Submitted (01-JUL-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
On Oct 16, 2003 this sequence version replaced gi:32362352.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 41693: contig of 41693 bp in length.
FEATURES
source
1. 41693
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OSJNBa0060M17"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 41693
Score: 4632.00 Matches: 949
Percent Similarity: 60.06% Conservative: 66
Best Local Similarity: 56.15% Mismatches: 55
Query Match: 80.17% Indels: 621
DB: 16 Gaps: 16
US-09-900-237A-30 (1-1080) x AC145384 (1-41693)
QY 7 AlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIleCysAlaAspGlyLeu 26
Db 11151 GCGAAGAAATCGGGGAAGCAA-----TGCCATGCTGCCAGATCTGCGGCGAGCGCGTG 11098
QY 27 GlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCysArgPheProValCys 46
Db 11097 GGCACGGCGGCGGACGGCGAGCTCTTCACCGCTCGGACGCTCGGGTCCCGGTGTGC 11038
QY 47 ArgProCysTyrGluHisGluArgGlyThrGlnAlaCysLeuGlnCysLysThr 66
Db 11037 CGGCCATGCTACGAGTACGAGCGCAAGGACGCGGCGGCTGCCGAGTGCAGGACC 10978
QY 67 LysTyrLysArgHisArg----- 72
Db 10977 AAGTACAGCGGCCCAA--GGGTATAAGTTGCACATTCATTCATCTCTCTTGGATT 10919
QY 72 ----- 72
Db 10918 TCACCTAAATACTTCATCTCTTTGGTTAAGAACAGCTACATTGGACATCTCGTGAAC 10859
QY 72 ----- 72
Db 10859 CCAATCAGGTTTGTGCATATTGTCATTTAGCAATAATAGTAGAAGATATATATGCT 10799
QY 72 ----- 72
Db 10798 TAGTGAATATGCTGAGTTAATAGCACAGAGTCTCTGTTCTTTAGTGAGCTGTATTGCAAGT 10739
QY 72 ----- 72
Db 10738 TGGTAGTGACAGTGACACATAAATTTCTTCTGTATGGGTATTATTCATTGAGAGAG 10679
QY 73 -----GlySerPro 75
Db 10678 CAAATGTGAGTGAATTTGTTGAGCTTAAACAAGATATTATCTTTGCTTCAGGAGCCG 10619

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
3 (bases 1 to 41693)
Buell,R.
Direct Submission
Submitted (16-OCT-2003) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
On Oct 16, 2003 this sequence version replaced gi:32362352.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 41693: contig of 41693 bp in length.
FEATURES
source
1. 41693
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OSJNBa0060M17"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 41693
Score: 4632.00 Matches: 949
Percent Similarity: 60.06% Conservative: 66
Best Local Similarity: 56.15% Mismatches: 55
Query Match: 80.17% Indels: 621
DB: 16 Gaps: 16
US-09-900-237A-30 (1-1080) x AC145384 (1-41693)
QY 7 AlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIleCysAlaAspGlyLeu 26
Db 11151 GCGAAGAAATCGGGGAAGCAA-----TGCCATGCTGCCAGATCTGCGGCGAGCGCGTG 11098
QY 27 GlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCysArgPheProValCys 46
Db 11097 GGCACGGCGGCGGACGGCGAGCTCTTCACCGCTCGGACGCTCGGGTCCCGGTGTGC 11038
QY 47 ArgProCysTyrGluHisGluArgGlyThrGlnAlaCysLeuGlnCysLysThr 66
Db 11037 CGGCCATGCTACGAGTACGAGCGCAAGGACGCGGCGGCTGCCGAGTGCAGGACC 10978
QY 67 LysTyrLysArgHisArg----- 72
Db 10977 AAGTACAGCGGCCCAA--GGGTATAAGTTGCACATTCATTCATCTCTCTTGGATT 10919
QY 72 ----- 72
Db 10918 TCACCTAAATACTTCATCTCTTTGGTTAAGAACAGCTACATTGGACATCTCGTGAAC 10859
QY 72 ----- 72
Db 10859 CCAATCAGGTTTGTGCATATTGTCATTTAGCAATAATAGTAGAAGATATATATGCT 10799
QY 72 ----- 72
Db 10798 TAGTGAATATGCTGAGTTAATAGCACAGAGTCTCTGTTCTTTAGTGAGCTGTATTGCAAGT 10739
QY 72 ----- 72
Db 10738 TGGTAGTGACAGTGACACATAAATTTCTTCTGTATGGGTATTATTCATTGAGAGAG 10679
QY 73 -----GlySerPro 75
Db 10678 CAAATGTGAGTGAATTTGTTGAGCTTAAACAAGATATTATCTTTGCTTCAGGAGCCG 10619

QY 76 AlaIleArgGlyGluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyr 95
Db 10618 CCGATCTTGGGATGAAGCGATGTTGATCGGATGATCTAGTGATGTAACAT 10559
QY 96 ProAlaSerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArg 115
Db 10558 CCAACATCCGCAACAGGACCAATGACACAGAGATGCCAGAGAGATGCTCACCTGGCGC 10499
QY 116 MetAsnThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGly 135
Db 10498 ATGAATCTCTGGAGGAATGATATATTGTCATCTCAATCTCAAGTATGATAGTGGTGGATGGT 10439
QY 136 LeuSerLysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSer 155
Db 10438 CATCCCAAGATGACAGTGTGAAATCCCTCGCATATATATCCATCGCTCACTCACAGC 10379
QY 156 Gln----- 156
Db 10378 CAGGTTTGTAGAAATTTTGTGATAGCGTGAAGTCGTTAACCTTCAGGGTATGTGCAA 10319
QY 157 -----MetSerGlyGluIle 161
Db 10318 AGATTGAACATAATATTAGTTCCTCTCGATTAAATCTATAGATCTCAGGTGAAT 10259
QY 162 ProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgAla 181
Db 10258 CTGGAGCGTCCCTCAT---CATATGATGTCTCCCGTTGGGAACATTTGGCAGACGTGG 10202
QY 182 ---ProPheProTyrValAsnHisSer----- 189
Db 10201 CATCCATTTCCCTATGTGAACCATTCACGTAAAGTTTATCTATCATCATCAAGATCTCATGTTT 10142
QY 190 -----ProAsnProSerArg 194
Db 10141 CGTTCATCGACTGGGAAATTTAATTTCTTTGTTGCTCAACAGCAACCATCAAG 10082
QY 195 GluPheSerGlySerIleGlyAsnValAlaTyrLysGluArgValAspGlyTyrLysMet 214
Db 10081 GAGTTTTCTGGTAGCTTGGCAATGTTGCATGAAAGAGAGATAGATGGCTGGAAATG 10022
QY 215 LysGlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGly 234
Db 10021 AAG--GATAAAGGTGCAATTTCCCATGGCTAATGTTACTAGCATTTGCTCTTCAAGAGT 9965
QY 235 ArgAlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeu--- 253
Db 9964 CGTGGGTTGGTGACATTTGATGATCTACTGACTATAACATGGAGGACGCTTACTGTGA 9905
QY 253 ----- 253
Db 9904 GTTCTGATATGCTTTGTCCACCGTTAGTCATCTATGAATCTTACCAGTAATTAATATTC 9845
QY 253 ----- 253
Db 9844 CTTGTATCTGGTGACTAACTTGCACATAGTTTGTACTTGGAGAACATAGGAAAAATCT 9785
QY 253 ----- 253
Db 9784 TGCATAATTCGAATATAATTAATTTTGTGTTTACAAGTATACCTTGTATTAAGCAAGGC 9725
QY 253 ----- 253
Db 9724 AATATGCACATTTAATGATCAGATTTAAATGGTTGTGATCTCTGAACCTTTATAATGC 9665
QY 254 -----AsnAspG1 256
Db 9664 GGACTTATCATATAGCTTGGCACTATATAATGTTGTCTATCTGTTATTACAGGAATGATGA 9605
QY 256 uThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrAr 276
Db 9604 AACAGCCAACTCTATCGAGAAAAGTGCCAAATTTGCTCATCCAGATAAATCCGTACAG 9545
QY 276 gMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAs 296

Db 9544 GATGGTTATTGTGCTCCGATGATGTCCTATGTAATATCTTGCACTACCGTATCACAAA 9485
QY 296 nProValArgAsnAlaTyrProLeuTyrLeuLeuSerValIleCysGluIleTyrPheAl 316
Db 9484 TCCTGTGCGTAATGATACACCGCTATGTTGCTCTCTGTCTATATGTCAGATTGGTTTC 9425
QY 316 aLeuSerTyrIleLeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLe 336
Db 9424 TCATCTCTGATTTCTGGATCAGTTCCTCCAAAGTGTCCCAATCAACCGTGAAACTTACCT 9365
QY 336 uAspArgLeuAla----- 340
Db 9364 GGATAGCTGGCTTTAAGGTGATGTACAATATTACCTCTGATACACCATTCCTCGCGTTA 9305
QY 341 -----LeuArgTyrAspAr 345
Db 9304 GCACCTAGTACTGATTAAGTACATAAATAATATCACTTGCACCTCTCAGGTATGACCG 9245
QY 345 gGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspProLe 365
Db 9244 GGAAGGTGAACCATCTCAATTTGGCTCTCTGTCATTTTGTCACTACTGTGATCTAT 9185
QY 365 uLysGluProProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrPr 385
Db 9184 GAAAGAACCTCTCTGTGTCACGCCAATACTGTGCTCTCATCTCTGTGTTGATACCC 9125
QY 385 oValAspLysValSerCysTyrValSerAspGlyAlaSerMetLeuThrPheAspAl 405
Db 9124 TCTTGATAAGGTATCTTGTCTGATGCTGATGAGCTGCAATCTTAACATTTGATGC 9065
QY 405 aLeuAlaGluThrSerGluPheAlaArgLysTyrValProPheValLysLysTyrAspIl 425
Db 9064 GTTCTGTGACCTTCAGAAATTTGTAGAAAATGGGTACCGTTCTGTAAAAAGTACAGCAT 9005
QY 425 eGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysVa 445
Db 9004 AGAACCCAGAGCTCCAGAGTGTACTTTCACAGAAAATTTGATTACTTGAAGACAAAAT 8945
QY 445 lGlnProSerPheValLysAspArgAlaMetLys----- 457
Db 8944 TCAGGCTTTCTTTTGTAAAGATCGTCTGCCATGAAGTAAATTTGGTTTATATGTGTCT 8885
QY 457 ----- 457
Db 8884 ACTTTTTTGTGTTTTTTTTTGTGTTAATCTCGCATTAATGCTGTTGCAAAACCAAAA 8825
QY 458 -----ArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuL 476
Db 8824 TCCAGAGGGAATATGAAGATTTAAAGTTCTGTGTTAATGCTCTCGTCCCAAGGCACAGA 8765
QY 476 ysValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrA 496
Db 8764 AAGTCTCTGAGAAAGGATGGAATATGCAAGATGGCACACCTTGGCTGTTAATAATACCA 8705
QY 496 rgAspHisProGlyMetIle----- 502
Db 8704 GAGACCATCTGGAAATGATTCAGGTAATATCATCTTAGATATGCGCATTTGGTGTGTT 8645
QY 503 -----GlnValPheLeuGlyHis 508
Db 8644 ATTGATCCAGCAATTTCTTACTCTTGACGCTCGTCATGCTCAGGTTTTCTTGGTCAT 8585
QY 509 SerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 528
Db 8584 AGTGGAGCCTTGATACCGAAGCAATGAGCTTCTCGTTTAGTCTATGTCTCTCGTAG 8525
QY 529 LysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAla-Leu----- 545
Db 8524 AAACGTCCTGGATTCCCAACCATAAAGAGCTGCTCTATGATGCACTTGTGAGTAGC 8465
QY 545 ----- 545

```
Db 8464 AGTAGAGTGGTGTGCTTAAGTCAGAGATATCTTGTTTAATTCATTACTTGATATGTG 8405
Qy 546 -----ValArgValSerAlaValLeuThrAsnGlyGlnIleMetLeuAs 560
Db 8404 GAATTTGAAACTGCAGGTTCGTGTATCAGCTGTCTTACTAATGAGCAATACTTGTGAA 8345
Qy 560 nLeuAspCysAspHisTyrIleAsnAsnSerIleAlaValArgGluAlaMetCysPheLe 580
Db 8344 TCTTGACTGTGATCACCATACATCAACATAGCAAGCTCTCCGAGAGGATATGTGCTTCT 8285
Qy 580 uMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGl 600
Db 8284 TATGGATCCAAACCTAGGAAGCGGTGCTGTATGTCCAAATTCCTCAGAGGTTTGACGG 8225
Qy 600 YIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAsp 616
Db 8224 TATCGATAGAAATGATCGATATCAAAACAGGAACACCGTGTGTTTTCGATGTAAAGCTTGGT 8165
Qy 616 ----- 616
Db 8164 TCCCTTTAAGATTCTTGGATGCTTCACTTCAGCTTTAGCAAAACATCAATAACTGTGTGAT 8105
Qy 616 ----- 616
Db 8104 GGATCAGCAATATTTATGTTTCTATGTGATGGAATTTGAACATTTCTGGAGTTTGTCAAGT 8045
Qy 617 -----IleAsnLeuArgGlyLeuAspGlyIleGlnGl 627
Db 8044 AACTTATCTGCATTTCTGCTGCTGCAGATTAATTTGAGAGGTCTCGATGTGCTCCAAGG 7985
Qy 627 yProValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluPr 647
Db 7984 ACCAGTTTATGTGGAAACCGGTGTGTGTTCAACAGAACAGCTCTTTATGGTTATGAAC 7925
Qy 647 oProIleLysAlaLysLysProGlyPheLeuAlaSerLeuCysGlyGlyLysLysLeuAl 667
Db 7924 CCCCATTAAAGCAGAGAGCCAGGTACTCTCTTCGCTTTGTGGGGACGAAAGAAC 7865
Qy 667 aSerLysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisValAspSerSe 687
Db 7864 AAAAAAGTCGAAGGAGAGAGACGCGGAAAGAAAAAGTACACACAAACATGTGGACAGTTC 7805
Qy 687 rValProValPheAsnLeuGluAspIleGluGlyValGlu 701
Db 7804 TGTGCCAGTTTTTAATCTTGAAGATATAGAGAGGGATTGAAGGTACCATCGAGTCTGG 7745
Qy 701 ----- 701
Db 7744 TTTTGTTCAAATTACATGTTTTGTCTCTAGATTTTCTAGATATTTTGGCCACCACCAT 7685
Qy 701 ----- 701
Db 7684 GCATTCGACCCCAACACTTTTTTTCGCAGTATTTCTGTATATGCTTAGGAAGGGAGGACA 7625
Qy 702 -----G1 702
Db 7624 CTAGGCTTAACCTTTCTCTTCACTATATATTTAGCTCGGCATCATTTTTTGTGGCAGG 7565
Qy 702 yAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPh 722
Db 7564 TTCTGGATTTGATGAGAAATCACTACTGATGCTCAATGAGCTTAGAGAAAGATT 7505
Qy 722 eGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValProGlnSe 742
Db 7504 TGGTCAATCTAGTGTGTTTTGTAGCTCCACTCTGATGAATATGGTGTGTTCTCAATC 7445
Qy 742 rSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAs 762
Db 7444 CGCAACTCCAGATCTCTCTGAAGAGGCCATACATGTTATCAGCTGTGGCTATGAAGA 7385
Qy 762 pLysSerGluTrpGlyThrGlu 769
Db 7384 CAAAAGTACTGGGAACTTGAGGTAAATAAGCTATGGCATTATGCTCTGCATTACATGC 7325
```

```
Qy 769 ----- 769
Db 7324 TACTAGCTTACCAGTACAAATTTACAATGTACAGAAATGTCGAGATATCAAGTGGTC 7265
Qy 769 ----- 769
Db 7264 TTTCTTGTACAAATTAAGGAGCAATTTGATTTCTGTTCTGATTCGACAAAGTACTTCT 7205
Qy 769 ----- 769
Db 7204 GAGTTCATGCGAAGAAACTTGACTAAATTTGGCATCTCTCTCAGTTGCCATTATAACTTT 7145
Qy 770 -----IleGlyTyrIleTyrGlySerValThrGluAspIl 781
Db 7144 GTTTATGATGCTATGCTACTTGCAGATTGGATGATCTATGGTTCTGTTACAGAAGATAT 7085
Qy 781 eLeuThrGlyPheLysMethHisAlaArgGlyTyrArgSerValTyrCysMetProLysAr 801
Db 7084 TCTCACTGGATTCAGATGTCATGCAGCTGGTGGCGGTCAATCTACTGCATGCCTTAAGCG 7025
Qy 801 gProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuAr 821
Db 7024 ACCAGCATTCAGGGGTCTGCCCAATTAATCTTTTCGATCGTCTGAACACGAGGTCTTCG 6965
Qy 821 gTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGlyTy 841
Db 6964 GTGGCTCTTGTTCTCTGGAAATTCCTTTTAGCAGGATTTGCCATATGGTATGGGTA 6905
Qy 841 rGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLe 861
Db 6904 CGGAGGACGGCTTAAGTTCTCTGGAGAGATTTGCTTTACATCAACACCACTATTATCCACT 6845
Qy 861 uThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGlyLy 881
Db 6844 CACATCGAATCCACTCTCTTGTACTGCATATGCGAGTATCTGCTCTCTCCTCCTCAGTGGAA 6785
Qy 881 sPheIleMetProGlu 886
Db 6784 GTTCATCATACAGAGGTAAATATGTAGAAATGTTTTTCTCTTCGACAGCAATCGC 6725
Qy 887 -----IleSerAsnLeuAlaSerIleT 894
Db 6724 AGACAGCATCGCTGACTGACTTATTTTGGACTGCAGATTAGCAACTTTGCGAGTATTT 6665
Qy 894 rPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerG 914
Db 6664 GGTATTATCTCTCTCTCTGTCATCTTTGTCTACTGGTATCCTTGAGATGAGTGGAGTG 6605
Qy 914 lYValGlyIleAspGluTyrTyrArgAsnGluGlnPheTyrValIleGlyGlyIleSerA 934
Db 6604 GTGTTGGCATCGATGATGTTGGAGGAATGAGCAGTTTGGGTTATTTGGAGTATTTCCG 6545
Qy 934 lAHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnP 954
Db 6544 CTCATCTATTTTCGGGTCTTCCAGGGTCTGCTGAAGGTGCTGCTGGAATGTATACACGCT 6485
Qy 954 hThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheL 974
Db 6484 TCACTGTCACTTCAAAAGGCTCTGATGAAGAAGGGGATTTTCGCTGAGCTCTCATGTGTC 6425
Qy 974 ysrTyrThrThrLeuLeuIleProThrThrIleLeuIleIleAsnMetValGlyValIV 994
Db 6424 AGTGGAACAACTTCTGATCCCAACGACCATCTATTTTGTATCATCAACCTGGTGGTGTG 6365
Qy 994 aLalaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyL 1014
Db 6364 TTGCTGGTATCTCTACGCTATCAACAGTGGCTACACAGTATGGGTCCGCTCTTTGGCA 6305
Qy 1014 ysrLeuPhePheAlaPheTyrValIleValHisLeuTyrProPheLeuLysGlyLeuMetG 1034
Db 6304 AGCTCTCTCTTCTGCTTCCAGCTGTATGTCACCTGTACCTCTTAAAGGGTCTCATGG 6245
```

QY 1034 lyArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIleP 1054
 Db 6244 GTGGCAGAACCGTACACCGACCATCTGTGTGTTTGGGCATCTCTCTTGCATCGATCT 6185
 QY 1054 heSerLeuLeuTrpValArgValAspProPheThrArgLeuAlaGlyProAsnIleG 1074
 Db 6184 TCTCTTGGCTGTGGTTCGTATCGATCATTCACACCGGTGTACAGCGCCGTGATACCC 6125
 QY 1074 lnThrCysGlyIleAsnCys 1080
 Db 6124 AAAATGTGTATCACTGC 6105

RESULT 12
 AC135958 154555 bp DNA linear PLN 25-APR-2003
 LOCUS Oryza sativa chromosome 3 BAC OSJNBa0059E14 genomic sequence,
 DEFINITION complete sequence.
 ACCESSION AC135958
 VERSION AC135958.2 GI:29837774
 KEYWORDS HTG.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 154555)
 Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Gansberger,K., Jones,K.M.,
 Overton II,L.L., Tsitrin,T., Kim,M.M., Bera,J.J., Jin,S.S.,
 Fadrosch,D.W., Tallon,I.J., Koo,H., Zisemann,V., Hsiao,J., Blunt,S.,
 Vanaken,S.S., Riedmuller,S.B., Utterback,T.T., Feldblyum,T.V.,
 Yang,Q.Q., Haas,B.J., Suh,B.B., Peterson,J.J., Quackenbush,J.,
 White,O., Salzberg,S.L. and Fraser,C.M.
 Oryza sativa chromosome 3 BAC OSJNBa0059E14 genomic sequence
 Unpublished
 2 (bases 1 to 154555)
 Buell,R.
 Direct Submission
 Submitted (01-NOV-2002) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 3 (bases 1 to 154555)
 Buell,R.
 Direct Submission
 Submitted (15-APR-2003) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 4 (bases 1 to 154555)
 Buell,R.
 Direct Submission
 Submitted (25-APR-2003) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
 On Apr 15, 2003 this sequence version replaced gi:24462343.
 Address all correspondence to:rice@tigr.org

BAC clone OSJNBa0059E14 is from Oryza sativa chromosome 3
 The orientation of the sequence is from Sp6 to T7 end of the BAC
 Clone.
 Genes were identified by a combination of several methods: Gene
 prediction programs including Fgenesh (<http://www.softberry.com/>),
 genscan and Genscan+ (Chris Burge,
<http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHM (Mark Borodovsky,
<http://genemark.biology.gatech.edu/GeneMark/>), and GeneSplicer
 (Mihaela Perteau and Steven Salzberg, contact mpertea@tigr.org),
 searches of the complete sequence against a peptide database and
 the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.shtml>).
 Annotated genes are named to indicate the level of evidence for
 their annotation. Genes with similarity to other proteins are named
 after the database hits. Genes without significant peptide
 similarity but with EST similarity are named as unknown proteins.
 Genes without protein or EST similarity, that are predicted by more
 than two gene prediction programs over most of their length are
 annotated as hypothetical proteins. Genes encoding tRNAs are
 predicted by tRNAscan-SE (Sean Eddy,
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are
 identified by repeatmasker (Arian Smit,

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

This BAC overlaps with rice BAC OSJNBb0015I02 (AC135563).

FEATURES
 source
 Location/Qualifiers
 1.154555
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="3"
 /map="near C393A"
 /clone="OSJNBa0059E14"
 /note="japonica cultivar-group"
 1373..1396
 /rpt_family="AT-rich"
 complement(2338..3123)
 /gene="OSJNBa0059E14.1"
 /note="EST AU165836, AU182644 from this gene"
 complement(join<2338..2570,2665..>3123))
 /gene="OSJNBa0059E14.1"
 complement(join(2459..2570,2665..3123))
 /gene="OSJNBa0059E14.1"
 /codon_start=1
 /product="unknown protein"
 /protein_id="AAP21417.1"
 /db_xref="GI:30103004"
 /translation="MAAIANLELSHTHTHTQVCSIDRPSRAHNGERSQPRGGG
 GURGVCGRGGGGRRGGVRRRERAGAVRAVPRAGGAAGSVLRREGAARD
 GDDGAARAVQVGAIVRAVRAPRAAPFGALQARTHPRRRRLHQDILLSSGG
 GDDGAAATREFVAIGA"
 2782..3013
 /rpt_family=" (CGCG)n"
 4169..5246
 /gene="OSJNBa0059E14.2"
 /note="similar to ubiquitin GB:BAC06474 GI:22004048
 (Clona savigny)"
 join(<4169..4410,4495..4975,5067..>5246)
 /gene="OSJNBa0059E14.2"
 join(4169..4410,4495..4975,5067..5246)
 /gene="OSJNBa0059E14.2"
 /codon_start=1
 /product="putative ubiquitin"
 /protein_id="AAP21418.1"
 /db_xref="GI:30103005"
 /translation="WRINVTLRGDRVALDVGGATTVAQVGMVAREIAVMORL
 FFAKCLDDHRLADYGVHDSVFLSURLATDAYQTEHNVRLMQPTATAKQEMH
 QQQQQLHVHVAADDEERAKRKPVSRRALKILSRQYDAWTQSHDAKFLDLLRHT
 GGGGARNVGELTGEDWSISRAELNAATGGFFVEQLQRRLGEFRFEAAARIKNHP
 RFSYDPRRVVAKQADWKNYILENPEAAAYEGRSPRHLGLRLAIFSGDGGGGGGGA
 KCRFTKARSLKRLNLFRLPKL"
 5138..5161
 /rpt_family=" (CGG)n"
 9301..11079
 /gene="OSJNBa0059E14.3"
 /note="contains Pfam profile PF03763 (remorin, C terminal
 region); EST BR041106.D41920.A0907433 from this gene"
 join(9301..10065,10558..11079)
 /gene="OSJNBa0059E14.3"
 join(9403..10065,10558..10749)
 /gene="OSJNBa0059E14.3"
 /codon_start=1
 /product="remorin-like protein"
 /protein_id="AAP21420.1"
 /db_xref="GI:30103007"
 /translation="MLHQHAPPPQPEPEVSLQSAATAADVAAGDDEEVTVVTTY
 RDIHLTPPTPTTTTTRGLGSAYSDTASSHSVSSEQFTMSREFTAMVAAGT
 TMOQTPDNGNGDQLTSTIGDELEETNPILAIVDPDHPHPIATPARSASOLEVVAAGP
 SPAPVEARQVKKEVEVTKVSAWQTAEVAKNNRFRREVINGWETQVEKASWLK
 KIERLDSQRAKALERTQNDIAKARKAEKASAEAKGLKLAKVLELANPMKAVGR
 VPTKRSFT"
 9422..9587
 /rpt_family=" (CCG)n"
 complement(12645..14126)
 repeat_region
 gene

mRNA	/gene="OSJNBa0059E14.4"			Best Local Similarity: 56.15%			Mismatches: 55		
	/note="similar to immediate early salicylate induced glucosyltransferase GB:AAB36653 GI:1685005 (Nicotiana tabacum)"			Query Match: 80.17%			Indels: 621		
	complement (<12645..>14126)			DB: 8			Gaps: 16		
	/gene="OSJNBa0059E14.4"								
CDS	complement (12645..14126)								
	/gene="OSJNBa0059E14.4"								
	/codon_start=1			US-09-900-237A-30 (1-1080) x AC135958 (1-154555)					
	/product="putative immediate-early salicylate-induced glucosyltransferase"								
repeat_region	/protein_id="AAP21423.1"								
	/db_xref="GI:30103010"								
	/translation="MAAATADGHRRLRVFFLPFFARGHLIPMTDLACLMRAASTD AVEVETAPVTPANAALATVAGNAARVVCVFPFVDVGLARGVECEGAAAHDTWV YRAVDLSRPAHEHLRRHPDAIVADYVFWATGVAEALGVPELRTFENPVGVFPOLAMN NVAVRDIIVREGADGPPVTPGMPGGREITIPVSELPDFLQDDHLVSMGWDRIKASQ LAGFVVNTPAALAEAPYCFEFSRVDARRAVFVGPSOPSGRAAAAARRGGDGDVDCI RWLSTKPSQVVVYCFSGWAHFSVTOTRELALGLEASNOPFLMVRSDSGDGGGERWE PEGWRMERGMVVRGWAPOLAVLAHPSVGAFVTHCGMNSVLEAARAGVPALTWPLV FPGFINERLVTEAAFGWEDGGKRGVAREAFVPGAVIRAVAGVFWAGGGGRR ERAAAMATALESARVAVGSSWRDIRLIQDLTASQP"								
	13319..13354								
repeat_region	/rpt_family="GC_rich"								
	13913..13963								
	/rpt_family="(CGG)n"								
	15663..21236								
gene	/gene="OSJNBa0059E14.5"								
	/note="highly similar to cellulose synthase catalytic subunit GB:AAC39336 GI:2827143 (Arabidopsis thaliana); EST_AU064228.AU176398.CB096725.AA751514 from this gene"								
	join(15663..16059,16193..16379,16712..16962,17065..17164,17243..17429,17725..17991,18085..18430,18519..18656,18739..18864,18950..19162,19322..19577,19773..19975,20219..20569,20652..>21236)								
	/gene="OSJNBa0059E14.5"								
CDS	join(16048..16059,16193..16379,16712..16962,17065..17164,17243..17429,17725..17991,18085..18430,18519..18656,18739..18864,18950..19162,19322..19577,19773..19975,20219..20569,20652..21236)								
	/gene="OSJNBa0059E14.5"								
	/codon_start=1								
	/product="putative cellulose synthase catalytic subunit"								
repeat_region	/protein_id="AAP21426.1"								
	/db_xref="GI:30103013"								
	/translation="MDGAKSGQCHVCQICGQGVGTAAADGELFTACDVCGFPVCRPCY EYERKDSQACQCKTKYKRHKGSPPIIGDESDDVADADSDVNYPTSGNDHKKHIA ERMLTWRMNSGNDIDVHSKYDSGEIPIYIPSLTHSQISGEIPGASP DHMSPVGNIGRGPFPYVNHSPNPSREFSGSLGNVAWKERVQDKMKDGAIPMAN GTSTAPSGRGVDIDASTDNMEDALLNDETRQPLSRKVPISSRINPYRMVILRL IVLCIFLHYRITNPVRNAYPLWLLSVCEIWFALSWILDQFPKSPINREYILDRAL RYDREGPSQLAPVDIFVSTVDPMKEPPLVANTVLSILAVIDPVKVKSVYDDGAA MLTFDALAESEFAKWPFCFKYSEIPEAPEWYPAQIDYLDKQVASFVKDRRAK REYEFKVRVNAVAKAQVPEBGMIMQDTPWPGNTRDHPMIQVFLHSGGLDTE GNEPLRVVLSREKPGFQHKKAGMALVRVSAVLTNGOYLLNDCDQIINNSKAL REAMCFMDPLNGRVVCYVQFPQRFQDGIIDRNDYANRNTVEFINLRGLDLOGGVYV RCGCFVNPLNGYEPPIKQKRPFPFSLCGRKKTKKSEKTEKKSHKRVDSVP VFNLEIDIEGSGFDEKSLMSLEKRFQSSVFVASTMEIGYGVPOSAIPES LLKEAHIVISCGVEDKSDWGTGIMWYGVSTEDILTGFMKHARGWRSIYCMKRPAPF GSAPINLDRNLQVRWALGSIIEILFSRHCRQIYWGGRKLFERFAVINTIYPLTS IPLLALCYLPCILLTGKFIPEISNFASIMFISLFSIFATGILEMWSGVGIDEW RNEQFWIGGISAHLPANFQGLLKVLAGIDTSFTVTSKASDEEGDFAELNFKWTLL IRPTTILINLVGVVAGISYAINSGWSGPLGKLFAPFVWIVVHLVFLPKGLMRQN RPTTIVVWMAILLASIFSLMLWRIDPFTTRVTGPDTPQKGINC"								
	gene	complement (21685..24525)							
/gene="OSJNBa0059E14.6"									
/note="similar to polygalacturonase precursor GB:AC26512 GI:3320462 (Cucumis melo)"									
Alignment Scores:									
Pred. No.:									
Score:									
Percent Similarity:									

```
QY 190 -----ProAsnProSerArg 194
Db 17197 CGTTGCATCGACTGGGAAATTTAAATTATTTCTTTGTTGCTCAACAGCAAAACCCATCAAGG 17256
QY 195 GluPheSerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMet 214
Db 17257 GAGTTTCTGGTAGCTTGGCAATGTTGTCATGGAAAGAGAGAGTAGATGCTGGAAATG 17316
QY 215 LysGlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGly 234
Db 17317 AAG---GATAAAGGTGCAATTCCTCGCTAAATGGTACTAGCATTCCTCTCAGAAGGT 17373
QY 235 ArgAlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeu--- 253
Db 17374 CGTGGGTTGGTGACATTTGATGATCTACTGACTATAACATGAGGAGCGCTTACTGTGA 17433
QY 253 ----- 253
Db 17434 GTTCTGATATGCTTTTGTCCACCGTTAGTCACTATGTAATTTACCAGTAATTAATATTC 17493
QY 253 ----- 253
Db 17494 CTTGTATCTGGTGTACTAACTTGCACTAGTTTGTACTTGTGAAGAACAATAGAAAAATTC 17553
QY 253 ----- 253
Db 17554 TGCATAATTCGAAATATAATTTTGTGTTTACAAGTATACCTTGTATTAGCAAGGC 17613
QY 253 ----- 253
Db 17614 AATATGCACATTTTAATGGATCAGATTTCTAAATGGTTGTACATCTCTGAACCTTTATAATGC 17673
QY 254 -----AsnAspG1 256
Db 17674 GGACTTATCATATAGCTTGCCACTATAAAGTTGTGCTATCTGTATTACAGGAATGATGA 17733
QY 256 uThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrAr 276
Db 17734 AACACGCCAACCTCTATCGAGAAAGTGCCAAATTCGTCTATCCAGATTAATCCGTACAG 17793
QY 276 gMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAs 296
Db 17794 GATGGTATTTGTCTCCGATTTGATTTGCTATGTATATTTCTTGCACTACCGTATACAAA 17853
QY 296 nProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAl 316
Db 17854 TCCTGTGCGTAATGCATACCCGCTATGGTTGCTCTCTGTCAATGAGATTGGTTGCTG 17913
QY 316 aLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrIle 336
Db 17914 TCTATCTTGGATTCTGGATCAGTTCCCAAGTGGTCCCAATCAACCCGTGAAACTTACCT 17973
QY 336 uAspArgLeuAla----- 340
Db 17974 GGATAGGCTGGCTTTAAGGTGATGTACAATAATTTACCTCTGATACACCAATTCCTCGTTA 18033
QY 341 -----LeuArgTyrAspAr 345
Db 18034 GCACGTAGTACTGATTACTGACAATAACAAAATATCATCTTGCATCTCAGGTATGACCG 18093
QY 345 gGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspProIle 365
Db 18094 GGAAGGTGAACCATCTCAATTTGGCTCTCTGTCACATTTTGTGACACTTTTGTGAGTGTGAT 18153
QY 365 uLysGluProProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrPr 385
Db 18154 GAAAGAACTCTCTTGTGACGTGCAATATCTGTCTCTCAATCTTGTGTTGGTATACCC 18213
QY 385 oValAspLysValSerCysTyrValSerAspGlyAlaSerMetLeuThrPheAspAl 405
Db 18214 TGTTCATAGGTATCTTGTATGTGTCTGATGAGCTGCAATGTTTAACTTTGATGC 18273
```

```
QY 405 aLeuAlaGluThrSerGluPheAlaArgLysTyrValProPheValLysLysTyrAspI1 425
Db 18274 GCITGTGTGACATTCAGAAATTTGCTAGAAATGGGTACCCGTTCTGTAAAAAGTACACAT 18333
QY 425 eGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysVa 445
Db 18334 AGAACCCAGAGCTCCAGAGTGGTACTTTGCACAGAAAATTTGATTACTTTGAAGACAAAGT 18393
QY 445 lGlnProSerPheValLysAspArgAlaMetLys----- 457
Db 18394 TCAGGCTTCTTTTGTAAAGATCGTCGTGCCATGAAGGTAAATTTGGTTTATATGTGTGT 18453
QY 457 ----- 457
Db 18454 ACTTTTGTGTGTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 18513
QY 458 -----ArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuL 476
Db 18514 TCCAGAGGGAATATGAAGAATTTTAAAGTTCTGTTAATGCTCTCGTCGCCCAAGGCACAGA 18573
QY 476 ysValPr-oGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrA 496
Db 18574 AAGTTCTCTGAGGAAGATGGAATTTATGCAAGATGGCACACCTTGGCCTGGTAAACAATACCA 18633
QY 496 rgAspHisProGlyMetIle----- 502
Db 18634 GAGACCATCTGGAATGATTTACAGTAAATATCACTTAGATATTTGCCATTTGTTGTTGTTG 18693
QY 503 -----GlnValPheLeuGlyHis 508
Db 18694 ATTGATCCAGCAATTTTCTTACTCTTGACGCTCTGGTCATGCTCAGGTTTCTTGTGTGTCAT 18753
QY 509 SerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 528
Db 18754 AGTGGAGGCTTGTATACCGAAGGCATGAGCTTCTCTGTTTGTAGTCTATGTCTCTGTGAG 18813
QY 529 LysArgProGlyPheGlnHisLysLysAlaGlyValaMetAsnAla-Leu----- 545
Db 18814 AAACGTCCTGGATTCCACACCATAAAAAGGCTGGTGTATGAATGCATTTGTGAGTAGC 18873
QY 545 ----- 545
Db 18874 AGCTAGAGTGGTGTGTTGTTAAGTAGAGAGATATCTTGTGTTTAAATTCATTACTTCATATG 18933
QY 546 -----ValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAs 560
Db 18934 GAATTTGAAACCTGCAGGTTCTGTTATCAGCTGTCTCTTACTAATGGAACAATCTTGTGAA 18993
QY 560 nLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLe 580
Db 18994 TCTTGACTGTGATCACTACATCAACATAGCAAAAGCTCTCCGAGAGGCTATGTGCTTCCT 19053
QY 580 uMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspG1 600
Db 19054 TATGATCCAAACCTTAGGAAGGCGTCTGTTATGTCCAATTTCCCTCAGAGGTTTGACCG 19113
QY 600 yIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAsp----- 616
Db 19114 TATCATGAAATGATGATATGCAATATGCAACAGGAACACCGGTGTTTTCGATGTAAGCTGTGT 19173
QY 616 ----- 616
Db 19174 TCCCTTTAAGATTTTGGATGCTTCACTTCACGCTTTAGCAACAATCAATACTGTTGAT 19233
QY 616 ----- 616
Db 19234 GGATGACCAATATTTATGTTTCTATGTGATGATTTGAACATTTCTGGAGTTTGTCAAAGT 19293
QY 617 -----IleAsnLeuArgGlyLeuAspGlyIleGlnG1 627
Db 19294 AACTTATCTGCATTTCTGCTGCTGCAGATTAATTTGAGAGGCTCGATGGTCTCCAAG 19353
QY 627 yProValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluPr 647
```

19354 ACCAGTTTATGTGGAAACCGGTTGTGTGTTCCACAGACACGCTCTTATGTTATGAACC 19413
647 oProIleLysAlaLysLysProGlyPheLeuAlaSerLeuCysGlyGlyLysLysAl 667
19414 CCCCATTAAGCAGAAGAGGCAGGTTACTTCTCTCGCTTTGTGGGGCAGCAAGAGAC 19473
667 aSerLysSerLysLysArgSerAspLysLysLysSerAsnLysHisValAspSerSe 687
19474 AAAAAAGTCGAAGAGAGACGCGAAAGAAAAAGTCAACAAACATGTGGACAGTTC 19533
687 rValProValPheAsnLeuGluAspIleGluGlyValGlu 701
19534 TGTCCAGTTTATCTTGAAGATATAGAGAGAGGATGAAGGTACCATCGAGTCTGG 19593
701 ----- 701
19594 TTTTGTTCATATTACATGTTTTTGTCTTAGATTTTGTAGTATTTTGGCCACCACCAT 19653
701 ----- 701
19654 GCATTGCACCCCAACAATTTTTTCGAGTATTTCTGTATATGCTTAGGAAGGGAGACA 19713
702 -----G1 702
19714 CTAGGCTTAACCTTTCTTACACTATATATTGACCTGGCACATCATTTTTGTGGCAGG 19773
702 yAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgph 722
19774 TTTCTGGATTTGATGATGAGAAATCACTACTATGCTCTCAATAGCTTAGAGAAAGATT 19833
722 eGlyGleSerAlaPheValAlaSerThrLeuMetGluThrGlyGlyValProGlnSe 742
19834 TGGTCAATCTAGTGTTTTGTAGCTCCACTCTGATGGAATATGTTGTCTTCAATC 19893
742 rSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGlyTyThrGluAs 762
19894 CGCAACTCCAGAACTCTCTCTGAAGAGGCCATACATGTTATCAGCTGTGGCTATGAAGA 19953
762 pLysSerGluTrpGlyThrGlu----- 769
19954 CAAAAGTGACTGGGGAACCTGAGGTAATAAAGCTATAGGCATTCATGCTCTGCATATACATGC 20013
769 ----- 769
20014 TACTAGCTACAGTACAAAATTATACAATGTAAACAGAAAATGTCAGATATCAAGTGTC 20073
769 ----- 769
20074 TTTCTTGTACAAAATTAAAGGACGCAATTTGATTTCTGTCTGATTCGCACAGTACTTCT 20133
769 ----- 769
20134 GAGTTTCATGCGAAGAAACTTGACTAAATGGCATCTCTCTCAGTTGGCAATTATAACTTT 20193
770 -----IleGlyThrIleTyThrGlySerValThrGluAspI 781
20194 GTTTATGATGCTPATGTGACTTCGCAATTGGATGGATCTATGTTCTGTTACAGAAGATAT 20253
781 eLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyCysMetProLysAr 801
20254 TCTCACTGGATTCAGATGTCATGCAGTGGCTGGCGGTCAATCTACTCATGCTCAAGG 20313
801 gProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuAr 821
20314 ACCAGCAITTCAGGGGTCTGCCCAATTAATCTTTCCGATCGTCTGAAACAGGTCGCTCG 20373
821 gTrpAlaLeuGlySerValGluIleuPheSerArgHisCysProLeuTyThrGlyTy 841
20374 GTGGGCTCTGTTCTGTGGAAATCTTTTATAGCAGGCAATCTCCCATATGATGAGGTA 20433
841 rGlyGlyArgLeuLysPheLeuGluArgPheAlaTyIleAsnThrThrIleTyProLe 861

20434 CGGAGGACGGCTTAAGTTCTCGGAGAGATTTGTTTACATCAACACCACCTATTATTCCACT 20493
QY uThrSerLeuProLeuLeuValTyCysIleLeuProAlaIleCysIleuLeuThrGlyTy 881
20494 CACATCATTCCTCTCTTGTACTGCATATGCCAGTATCTGTCTCTCCTCCTCCTCCTCCT 20553
QY sPheIleMetProGlu----- 886
20554 GTTCATCATACACAGAGGTAATATGTAGATTTGTTTTCCTCTCTTCGACAGAAATCGC 20613
QY -----IleSerAsnLeuAlaSerIleT 894
20614 AGAAGCATCGCTGACTGACTTATTTTGGACTGCAGATTAGCAACTTTCGAGTATTT 20673
QY rpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerG 914
20674 GGTTTATCTCTCTCTTCTGTCAATCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 20733
QY lyValGlyIleAspGluTrpArgAsnGluGlnPheTrpValIleGlyIleSerA 934
20734 GTGTTGGCATCGATGAATGTTGGAGGATGAGCAGTTTGGGTTATTGGAGGTTATTCG 20793
QY laHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn 954
20794 CTCATCTATTTGGCTCTTCCAGGCTCTGCTGAAGTCTGCTGCTGCTGCTGCTGCTGCTG 20853
QY heThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyMetPheL 974
20854 TCCTGTCACTCAAGGCTCTGATGAAGAGGCGATTTGCTGAGCTCTACATGTTCA 20913
QY ystTrpThrLeuLeuIleProThrThrIleLeuIleIleAsnMetValGlyValV 994
20914 AGTGGCAACCTCTCTGATCCACCGACCATTTTGTGATCATCAACCTGCTGCTGTTG 20973
QY alAlaGlyThrSerTyAlaIleAsnSerGlyTyThrGlnSerTrpGlyProLeuPheGlyL 1014
20974 TTGCTGTGATCTCTCTGCTATCAACAGTGGCTACAGTCATCGGCTCGCTCTTTGGGA 21033
QY ysLeuPhePheAlaPheTrpValIleValHisLeuTyThrPheLeuLysGlyLeuMetG 1034
21034 AGCTCTCTCTTGTGCTTCTGCTGATTTGTCACCTGTACCCCTTCTTAAAGGCTCTCATGG 21093
QY lyArgGlnAsnArgThrProThrIleValIleValIleValIleValIleValIleValIle 1054
21094 GTGGCAGAACGTCACCGACCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 21153
QY heSerLeuLeuTrpValArgValAspPropheThrThrArgLeuAlaGlyProAsnIleG 1074
21154 TCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 21213
QY lnThrCysGlyIleAsnCys 1080
21214 AAAATGTTGTTATCAACTGC 21233
RESULT 13
AF150630 3723 bp mRNA linear PLN 04-FEB-2000
LOCUS Gossypium hirsutum cellulose synthase catalytic subunit (celA3)
DEFINITION mRNA, complete cds.
ACCESSION AF150630 AF200453
VERSION AF150630.2 GI:6446576
KEYWORDS Gossypium hirsutum (upland cotton)
SOURCE Gossypium hirsutum
ORGANISM Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 3723)
Laosinchai,W., Cui,X. and Brown,R.M. Jr.
A full length cDNA of cotton cellulose synthase has high homology
with the Arabidopsis RSW1 gene and the cotton CelA1 gene (Accession
No. AF200453) (PCR 00-002)
JOURNAL Plant Physiol. 122 (1), 291 (2000)

REFERENCE 2 (bases 1 to 3723)
AUTHORS Kimura,S., Laosinchai,W., Itoh,T., Cui,X. and Brown,R.M. Jr.
TITLE Immunogold Labeling of Rosette Terminal Cellulose Synthesizing Complexes in a Vascular Plant (Vigna angularis)
JOURNAL Unpublished
REFERENCE 3 (bases 1366 to 3470)
AUTHORS Laosinchai,W. and Brown,R.M. Jr.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-1999) Botany, University of Texas at Austin, W. 24th Street, Austin, TX 78713, USA
REFERENCE 4 (bases 1 to 3723)
AUTHORS Laosinchai,W., Cui,X. and Brown,R.M. Jr.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-1999) Molecular Genetics and Microbiology, University of Texas at Austin, Austin, TX 78712, USA
REMARK Sequence update by submitter
COMMENT On Nov 17, 1999 this sequence version replaced gi:5081778.
FEATURES
Location/Qualifiers
1..3723
/organism="Gossypium hirsutum"
/mol_type="mRNA"
3 /cultivar="Texas marker-1"
/db_xref="taxon:3635"
/tissue_type="cotton fiber"
/dev_stage="24 days post anthesis"
1..3723
/gene="celA3"
289..3472
/genes="celA3"
/note="celA1 homolog"
/codon_start=1
/product="cellulose synthase catalytic subunit"
/protein_id="AAD39534.2"
/db_xref="GI:6446577"
/translations="MESEGDIGKPKMNLGGQTQCICGNVGNKNTDGDFFIACNICAF
PVCRCYERKDNQSCPCQKTRYKWKQSGFALLGDRGTGGDADDGASDFIYSENQ
QKQKLAERNQWNAKTYGRGDEGAPTYDKEISHNHIPLLTSGQVSGEISAASPERLS
MASPVAGKSSIRVVDVPRVREFSGGLGNVAMKRVDMKQKNTVPMSTCOATSE
RGLGDDIDSTDLVDDQLNDARQPLSRKVSVSSKINPYRMVILRLVILICFLHY
RTNPVPNAVALLVLSVCEIWFATSWILDQPKWLPVNRIFYLRLALRYDREGPS
ELAAVDIFVSTVDPLKEPFLVNTVLSLAVDYPVKVSCVSDDDGAAMLTFEALSE
TSEFAKQWFFCKKNIEIPRAPEWYFAQKIDYKQVTSFVKDRRAMREYEEFKVR
INGLAKAQKPEEGIMQDTPPGNNTDHPGMIOVFLGSGGLDARGLNELPRIVY
VSREKRPFGKHKAGAMNALRVSAVLNNGAFLLNLDCHDHYINNSKALUREAMFLMD
PNLHGQCVVQPPORFDGIDRNDRYANRNVFEDINRLGDLGQGPVYVGTGCVENRT
ALYGVPEPLPKPKHRTGIISSLGSGSRKSSKSKGSKKSGKHVDSTVPVFNLED
IEBGEVAGFDEKSLIMSQMSLERFGQSAVFVASTLMENGVFQSPATPEPILKEAI
HVISCYEDKTWMSGEIWIYGSVTEIDILTFKMHARGWRSIYCMKPRPAPKGSAPIN
LSDRLNQLRWALGSVEILFSRCHPIWYSGRLKWLRFAYVNTIYPTVAIPILMY
CTPAPVCLLTNKFIIPOISNLAIWFIISLFLSIFATILKMKWNGVIGIDQWNRNBOFW
VTGGVSAHLFAVFGKLLKVLGIDNTFTVTSKASDEDDGDFAEIYMFKMTLLIIPPTLL
LIINLVGVVAGISYVINSQYSGWGLFGKLFPFAMVILHLFPFLKGLMGRQRTTIV
VWMSILLASIPSLWLVRIDPFTTRVTGPDVEQGGINC"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3723
Score: 4577.50 Matches: 853
Percent Similarity: 87.83% Conservative: 100
Best Local Similarity: 78.62% Mismatches: 105
Query Match: 79.22% Indels: 27
DB: 8 Gaps: 11
US-09-900-237A-30 (1-1080) x AF150630 (1-3723)

QY 2 AspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIle 21
Db 278 GAAGGAGATATTGGGGGC---AAGCCCATCAAGACATCTTGGTGGCCAGACATGCCAAATC 334
QY 22 CysAlaAspGlyLeuGlyThrLeuAspGlyAspValPheThrAlaCysAspValCys 41
Db 335 TGTGTGACAAATGTGGCAAAATACAGATGATGATCCATTCATTGCTTGCAATATAGT 394

QY 42 ArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAlaCys 61
Db 395 GCATTTCTCTGTATGCGAGGCTTGTATGATATGAAGAAGAGGACGCGGAATCAGTCTTGC 454
QY 62 LeuGlnCysLysThrLysTyrTyrArgHisArgGlySerProAlaIleArgGlyGlu 81
Db 455 CCTCAATGCAAGACACAGATACAAAGTGCAGAAAGGAGCAGTCTCTATTCTTGAGACAGG 514
QY 82 GlyAspAspThrAspAlaAspGly---SerAspPheAsnTyrProAlaSerGlyThr 100
Db 515 GAAACGGGTGGTGCATGATGATGCGCAGTGATTTTCAATTTAC-----TCAGAAAAC 568
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
Db 569 CAAGACACAGACAGAAACTTGCAGAAAGATGCGGGCTGGGAATCAAAGTATGTAGTA 628
QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 629 GGGGAGGATGTTGGTGTCTCCAACCTTATGATAA----- 661
QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln---MetSerGly 159
Db 662 -----GAAATTTCTCATAAACCATATCTCTCTCACCAGTGGCGGAGGAGTTCCTGGT 715
QY 160 GluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArg 179
Db 716 GAGCTATCTGCTGCATCCTGAGCGGCTTCTATGGCATCTCTCGAGATTGCTGTGGTGG 775
QY 180 ArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe---SerGly 198
Db 776 AATCGAGTATTAGGTTCTGTG-----GATCCAGTGAGGAGTTGGTTCATCA 823
QY 199 SerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLys 218
Db 824 GGACTAGGCAATGATAGCTGGAGGAGAGATTGATGGCTGGGAAGATGAAGCAGGAGAAG 863
QY 219 GlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThr 238
Db 884 AATACGTGTTCTCTATGAGT-----ACATGCCAGGTACTTCTGAA---AGAGGCTCTTGA 934
QY 239 AspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArg 258
Db 935 GATATTGATGCCAGCACTGATGCTCGTGGATGACTCTCAACTGAATGATGAAGCTAGA 994
QY 259 GlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetVal 278
Db 995 CAGCCTCTATCAAGGAAGGTTTCCGTTTCTCTATCAATAATAATCTTATCGGATGGTC 1054
QY 279 IleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProVal 298
Db 1055 ATCACTACGCGTCTGTATTCTCTGCAATTTCTTGCACTACCGAATAACAACCCCTGTT 1114
QY 299 ArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSer 318
Db 1115 CCAAAATGCCTATGCTATGTTGATATCTGTGATATGAGATTGGTTGGTTCATATCC 1174
QY 319 TrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArg 338
Db 1175 TGGATATTGGATCAGTTCCTCCAAATGGCTTCTCTGTTAAATCGTGAGACGATCTTGTAGAG 1234
QY 339 LeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePhe 358
Db 1235 CTTGCTTAAAGATATGATCGAGAGGAGAACCGTCGGAATTAGTGCAGTTGCATATTT 1294
QY 359 ValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSer 378
Db 1295 GTGAGTACTGTCCGCCCATTAAGAGAGCGCTCCACTTGTGACAGCCAACTGCTACTATCC 1354
QY 379 IleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspGlyAla 398
Db 1355 ATTCCTGCTGTTGACTACCCAGTTCACAAAGGCTCTCTTGCTATGTTTCTGATGATGGTGT 1414
QY 399 SerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValPro 418

Db 1415 GCCATGTTTAACTTTGAGCACTGCTGAGCATCAGAGTTGCCGGAATGGTACCT 1474
Qy PheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIle 438
Db 1475 TTCTGCAAGAAATATAACATTGAACCTCGGGCTCCAGAAATGGTATTTTGCACAGAAGATT 1534
Qy AspTyrIleLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArg 458
Db 1535 GACTACCTGAGGATAAGTTCAACACTCTTTTGTAAAGATCGTAGAGCTATGAAGAGA 1594
Qy GluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValPro 478
Db 1595 GAATATGAGAGATTAAAGTTTCGATTAATGCGCTTGTCCGAAGGCACAGAAATTCCT 1654
Qy GluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHis 498
Db 1655 GAAGAAGATGGATTATGCGAGATGGTACACCATGGCTGGTGAATTAACACCCAGAGACCAT 1714
Qy ProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGlu 518
Db 1715 CCAGGAATGATCCAGGTTTCTTAGGCCAGAGTGGGGCCCTTGATGCTGAGGGCAATGAG 1774
Qy LeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLys 538
Db 1775 CTGCCAAGTTAGTTTATGTTCTCGTGAAGCGTCCAGGCTCCACACACACAAAAA 1834
Qy AlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMet 558
Db 1835 GCTGGTGTATGAATGCACCTGTTTCGAGTGTGAGTGTCCAGTCTTACTTAATGAGGCTTCTGTG 1894
Qy LeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCys 578
Db 1895 TTAAATCTTGATGTGATCACTACATAAATATGCAAGGCTTTAAGGGAAGCTATGTGT 1954
Qy PheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPhe 598
Db 1955 TTCTGATGGATCCAACTTGGAAAGCAAGTTGTTACGTTCACTTCTCTCAAGGTTT 2014
Qy AspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsn 618
Db 2015 GATGGTATTGATCGAATGATCGATATGCCAACAGAAATACGTGTTTCTTGTATATAAC 2074
Qy LeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsn 638
Db 2075 TTGAGGGCTCTGGATGGCAATTCAGGGCCGGTTTATGTCGGTACGGATGTGTTTCAAT 2134
Qy ArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLys-----LysProGlyPhe 656
Db 2135 AGAACAGCATTTGATGGATATGAACCCCTTAAGCCCAAGCATAGAAAAACAGGATA 2194
Qy LeuAlaSerLeuCysGlyGly---LysLysLysAlaSerLysSerLysLysArgSerSer 675
Db 2195 CTATCTTCAATTGTCGGAGGTTCTGTAAGAGAGATTCAAAATCAAGTAAAGGATCA 2254
Qy AspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeuGluAsp 695
Db 2255 GACAGAAGAATCTGGCAAGCATGTTGATTTCTACTGTACAGATTAATCAATCTAGAGAT 2314
Qy IleGluGluGlyValGluGlyValaGlyPheAspAspGluLysSerValLeuMetSerGln 715
Db 2315 ATAGNAGGGGTGTGAAGGGCTGGGTTTGATGATGAAGTCAATTAATCTAGATGTCAAA 2374
Qy MetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGlu 735
Db 2375 ATGAGCCTGGAGAAACGGTTTGGCCAGTCTGCTGTTTGTGCTCTACACTTATGAA 2434
Qy TyrGlyGlyValProGlnSerSerThrProGluSerLeuLysGluAlaIleHisVal 755
Db 2435 AATGTGTGTCTCCCTCAGTCTGCTACCCCGGAACACTTCTTAAGGAAGCAATTCATGTT 2494
Qy IleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGly 775

Db 2495 ATCAGCTGTGGATACGAAGACAAAAACAGACTGGGGAAGTGAGATTGGGTGGATCTATGTT 2554
Qy SerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrAspSerVal 795
Db 2555 TTCTGTACAGAAGATATCTTACTTGGATTCAAAATGATGCCCGTGGTGTGGCGTCAATT 2614
Qy TyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArg 815
Db 2615 TACTGCATGCCMAACGCCCCAGCTTTAAAGGTTCTGCTCTATCAATCTTTCAGATCGT 2674
Qy LeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluLeuLeuPheSerArgHisCys 835
Db 2675 CTGAACCAAGTCTCGATGGCTCTTGGTCTCTGGAAATCTCTTTCAGTAGGCAATGTC 2734
Qy ProLeuTyrTyrGlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsn 855
Db 2735 CCTATCTGGTATGGTTATAGTGGTAGCTCAAGTGGCTTGAACGATTTCGCATATGTTAAC 2794
Qy ThrThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIle 875
Db 2795 ACAACCATCTACCCAGTCACCTGCTATTCCTCTCTCATGTACTGCACACTGCCAGCAGTC 2854
Qy CysLeuLeuThrGlyLysPheIleMetProGluLeuSerAsnLeuAlaSerIleTyrPhe 895
Db 2855 TGTCTGCTCACTAAACAAGTTCATTATTCCACAGATTAGTAACCTTGCAGTATATATGTTT 2914
Qy IleAlaIlePheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyVal 915
Db 2915 ATATCCCTCTTCTCTCATCTTTTGCACCTGGTATTTTAAAAATGAAGTGAATGGTGTG 2974
Qy GlyIleAspGluTyrTyrArgAsnGluGlnPheTyrValIleGlyGlyIleSerAlaHis 935
Db 2975 GGAATTCACCAATCGTGGAGAAATGAACAGTTTGGGTATCGGTGCTGCGGCACAT 3034
Qy LeuPheAlaValPheGlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThr 955
Db 3035 CTTTTCGTGTTTCCAAAGGCTTCTCAAAGTTCCTGTCGCAATTGACACTTAACCTTCACG 3094
Qy ValThrSerLysAlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTyr 975
Db 3095 GTCACTTCAAAAGCATCAGACGAAGATGGGATTTTGTGCACTTTACATGTTCAAAATGG 3154
Qy ThrThrLeuLeuLeuProProThrThrIleLeuIleLeuAsnMetValGlyValValAla 995
Db 3155 ACACCCCTTCTCATCCACCAACCACTCTCTCATATAAATCTGTCGGGTGTTGTGCA 3214
Qy GlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeu 1015
Db 3215 GGAATATCGTATGTATCAATAGTGGCTACCAATCATGGGTGCCCTCTTTGGTAAAGCTA 3274
Qy PhePheAlaPheTyrValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyValArg 1035
Db 3275 TCTTTGCTCTCTGGTGATCATCATCTCTACCCCTTCTTGAAGAGTTTGAAGGAGCGC 3334
Qy GlnAsnArgThrProThrIleValIleValThrAlaValLeuLeuAlaSerIlePheSer 1055
Db 3335 CAAACCGGTACACCTACCATTTGTTGTATGGTCAATTCGTCTGCTCCATTTTCTCT 3394
Qy LeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThr 1075
Db 3395 TTGTTGGTGGCGGATCGATCCCTTCAACACAGAGTTACCGGGCCGAGTGTAGAACAG 3454
Qy CysGlyIleAsnCys 1080
Db 3455 TGTGGAATAAACTGT 3469
RESULT 14
LOCUS BT002335 3229 bp mRNA linear PLN 15-DEC-2002
DEFINITION Arabidopsis thaliana clone C104938 putative cellulose synthase
catalytic subunit (At5g05170) mRNA, complete cds.
ACCESSION BT002335
VERSION BT002335.1 GI:26983831

FLI CDNA.	US-09-900-237A-30 (1-1080) x BT002335 (1-3229)
Arabidopsis thaliana (thale cress)	
Arabidopsis thaliana	
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsids.	
REFERENCE	
1 (bases 1 to 3229)	
AUTHORS	
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.	
Arabidopsis Open Reading Frame (ORF) Clones	
Unpublished	
2 (bases 1 to 3229)	
AUTHORS	
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.	
Direct Submission	
Submitted (15-DEC-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA	
Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.	
COMMENT	
FEATURES	
source	
1..3229	Location/Qualifiers
/organism="Arabidopsis thaliana"	
/mol_type="mRNA"	
/db_xref="taxon:3702"	
/chromosome="5"	
/clone="C104938"	
/ecotype="Columbia"	
/note="This clone is in PUNI 51."	
gene	1..3229
/gene="At5g05170"	
CDS	1..3198
/gene="At5g05170"	
/codon_start=1	
/evidence=experimental	
/product="putative cellulose synthase catalytic subunit"	
/protein_id="AA086168.1"	
/db_xref="GI:26983832"	
/translation="MESEGTAGKPMKNIVPOTCOICSDNVGKTVGDRFVACDICSF PYRCPEYERKDNQSCPOCKRVKRLKGSFAIPGDKDEDGLADEGTVERFNPQKEK ISERMLGHLTKGKEGEPQYDEKVEVSHNHLPLRTSQDTSGEFSAASPERLSYST IAGKRLPYSSDVNGSPNRRIVDPVGLGNVANKERDGMKQKQKNTGTEPVSQAASER GGVDIDASTDILLADALLNDEARQPLSRKVSIPSSRIINPYRMVIMLRVILCLFLHYR ITNPVNFALWLSVCIETWFAISWILDQEPFVFNRETYLDRALRYDREGPSQ LAAVDIFVSTVDPLPEPLVTANTVLSILAVDPVKVSCYVSDGAAMLSPELSAET SEFARKWVPFCCKYKIEPRAPWYFAAKIDYLDKQVTSFVKDRAMKREYEEFKIRI NALVSKALKCEEGWMDGTWPNGNTRDHPGMQVFLGONGGLDAGNELPRLVV SREKPGQHHKAKAMALVRVSAVLNPGFLNLDCHYINNSKALREACMFLLMDP NLGQVCVQFPQREDDGIDKNRYANRNTVPFDINLRGLDGQGPVYVGTGCVFNRTA LYGEPPFKVHKHLSLSLKGCGSKNSKAKESDKKSGRHTSDTVPVFNLDIE EGVEGADKXKALMSQLEKRFQSAFVASTLMENGVPVPSATPENILKEAIIHV ISCGVEDKSDMEIGTWYGVTEIDLTGFKMHARGWSIYCMPLKPAFKGAPNLNS DLNOLVRLWALGSVILFSRHCPFIWYGNLKLFLERFAYVNTTVIPIITSIPLMYCT LPAYCLFNQPIIPQISNIASINFLSFLSIPATGILEMRWSGVIDSWRNBNQFVVI GPVSHLFAVQGLDKLUGADITNFTVTSKASDEGDFAEFLYFLKMTLLIPPTILLI VNLGVAGVSGYAINSGYQSGWPLFGKLFAPFVIVHLVFLKGLMGRQNRTPITVVV WSVLLASTFSLIWLVRIDPFTSRVTGPDILECGING"	
3' UTR	3199..3229
/gene="At5g05170"	
ORIGIN	
Alignment Scores:	
Pred. No.:	0
Score:	4530.00
Percent Similarity:	87.82%
Best Local Similarity:	78.63%
Query Match:	78.40%
DB:	8
Length:	3229
Matches:	839
Conservative:	98
Mismatches:	104
Indels:	26
Gaps:	8
QY	19 CysGlnIleCysAlaAspGlyLeuGlyThrLeuAspGlyAspValPheThrAlaCys 38
DB	58 TGCAGATCTGTAGTGACAAATGTTGGCAAGACTGTTGATGGAGATCGTTTGTGGCTTGT 117
QY	39 AspValCysArgPheProValCysArgProCysTyrGluHisCysArgGlyGluGlyThr 58
DB	118 GATATTGTTTCATTCCAGTTTGTTCGGCCTTGTACGAGTATGAGAGGAAGATGGGAAT 177
QY	59 GlnAlaCysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArg 78
DB	178 CAACTCTGTCTCAGTCGAAACACAGATACAGAGGCTCAAGGTAGTCTCTGATATCCT 237
QY	79 GlyGluGlyAspAspThrAspAlaAspGlySer---AspPheAsnTyrProAla 97
DB	238 GGTGATAAAGACGAGGATGGCTTAGCTGATGAAGGTACTGTTGAGTTCAACTACCCCT--- 294
QY	98 SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsn 117
DB	295 -----CAGAAGGAGAAAATTCAGAGCGGATGCTTGGTTGGCATCTTACT 339
QY	118 ThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSer 137
DB	340 CGTGGGAAGGAGAGGAATGGGGAACCCCACTATGATAA----- 381
QY	138 LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln--- 156
DB	382 -----GAGGTCTCTCACATCATCTTCTCGTCTCAGCAGCAGACACAAGAT 426
QY	157 MetSerGlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsn 176
DB	427 ACTTCAGGAGAGTTTCTGCTGCCTCACCTGAACGCCTCTCTGTATCTTCTACTATCGGT 486
QY	177 IleSerArgArgAlaProPheProTyr---ValAsnHisSerProAsnProSerArgGlu 195
DB	487 GGGGGAAGCGCTTCCCTATTTCATCAGATGTCATCAATCACCATAAGAGGATTGTG 546
QY	196 PheSerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLys 215
DB	547 GATCCTGTTGGAGTCTGGGAATAGCTTGGAAAGAGAGAGTTCATGGCTGGAAAAAGAA 606
QY	216 GlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg 235
DB	607 CAAGAGAAGAACTACTGTCCTGTC-----AGCACCGACGCTGCTTCTGAA---AGA 654
QY	236 AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp 255
DB	655 GGTGGAGTAGATATTGATGCCAGCACAGATATCTCAGCAGATGAGGCTCTGCTGAATGAC 714
QY	256 GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr 275
DB	715 GAAGCGAGCGACGCTCTGTCAAGGAAAGTTTCAATTCCTTCTATCAGCGATCAATCCTTAC 774
QY	276 ArgMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThr 295
DB	775 AGAATGGTTATTATCTCGCGGCTTTGTTATCCTTTCTCTCTTTCATCCGATTAACA 834
QY	296 AsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPhe 315
DB	835 AACCCAGTGCCAAATAGCCTTCTCTATGGCTGGTCTCTGTGATATGTGAGATCTGGTTT 894
QY	316 AlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyr 335
DB	895 GCCTTATCTCTGGATTTCGATCAGTTTCCCAGTGGTTTCTCTGTAACCGTGAACCCCTAC 954
QY	336 LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaVal 355
DB	955 CTCGACAGCCTTGCCTTAAGATATGATCGTGNAGGTGAGCCATCAGTACGTCTGCTGT 1014
QY	356 AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr 375

Db 1015 GACATTTTGTGAGTACTGTTTGACCCCTTGAAGGAGCCACCCCTTGTGACAGCCACACACA 1074

Qy 376 ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp 395

Db 1075 GTGCTCTCTATCTGGCTGTTGACTACCCAGTTGACCAAGGTGCTCTGTATGTTCTGAT 1134

Qy 396 AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys 415

Db 1135 GATGGTGCTGCTGATGTTATCATTTGAATCACTTCAGAAACATCAGAGTTTGCTCGTAAA 1194

Qy 416 TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys 435

Db 1195 TGGGTACCATTTTGCAGAAATATAGCATAGAGCCTCGTGACACAGAAATGGTACTTTGCT 1254

Qy 436 GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla 455

Db 1255 GCGAAAAATAGATTACTTTGAAGGATAAAGTTTCAGACATCATTTGTCAAAGATCGTAGAGCT 1314

Qy 456 MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu 475

Db 1315 ATGAAGAGGGATATAGAGAAATTTAAATCCGAATCAATGCACATTGTTTCCAAAGCCCTA 1374

Qy 476 LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThr 495

Db 1375 AAATGTCCTGAAGAAGGTGGTATTATCAAGATGGCACACCGTGGCCTGGAAATAATACA 1434

Qy 496 ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu 515

Db 1435 AGGACCATCCAGGAATGATCCAGTCTTCTTAGGGCAAAATGGTGGCATTTGATGCGAG 1494

Qy 516 GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis 535

Db 1495 GGCNAATGAGCTCCCGCTTGGTATATGTTTCTCGAANAAGCCACAGGATTCAGGCAC 1554

Qy 536 HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly 555

Db 1555 CACAAAAGGCTGGTGCTATGAATGCATCGTGAGAGTTTCAGCAGATTCCTTACCNAATGGA 1614

Qy 556 GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu 575

Db 1615 CCTTTCATCTTGAATCTTGAATGTGATCATTTACATAATAACAGCAAGCCCTTAAGAAA 1674

Qy 576 AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro 595

Db 1675 GCAATGTCTTCCATGGAGCCCAAACTCGGAAGCAAGTTTGTATGTTTCAGTTCCCA 1734

Qy 596 GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe 615

Db 1735 CAAAGATTTGATGTTATCGATAAGAACGATAGATATGCTAATCGTAATACCGTGTCTTCT 1794

Qy 616 AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys 635

Db 1795 GATATTAACTTGAGAGGTTTATAGTGGGATTCAGGACCTGTATATGTGCGAACTGGATGT 1854

Qy 636 ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAla-----LysLys 653

Db 1855 GTTTTCAACAGACAGCATATACGGTTATGACCTCCAAATAAAAGTAAACACAGAGAG 1914

Qy 654 ProGlyPheLeuAlaSerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArg 673

Db 1915 CCAAGTCTTTTATCTAAGCTCTGTGGTGATCAAGAAAGAAAGAAATTCAAAAGCTAAGAAA 1974

Qy 674 SerSerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu 693

Db 1975 GAGTCGGCAAAAGAAATACAGGAGCATCTGACTCAACTGTTTCTCTGTATTCAACCTC 2034

Qy 694 GluAspIleGluGluGlyValGluGluValaGlyPheAspAspGluLysSerValLeuMet 713

Db 2035 GATGACATAGAAAGGGAGTTGAAGGTGCTGGTTTGTATGATGAGAAAGCGGCTCTTAATG 2094

Qy 714 SerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeu 733

Db 2095 TCGCAAAATGAGCCTGGAGAGCGGATTTGGACAGTCTGCTGTTTGTGTTGCTTCTACCCCTA 2154

Qy 734 MetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIle 753

Db 2155 ATGAAATAGTGGTGTCTCTCTCAGCAACTCCAGAAACCTTCTCAAGAGGCTATC 2214

Qy 754 HisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIle 773

Db 2215 CATGTCAATTAGTTGTGTTATGAGGATAAGTCAGATTGGGGAATGGAGATTGGATGATC 2274

Qy 774 TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArg 793

Db 2275 TATGGTTCTGTGACAGAGATATTCTGACTGGGTTCAAAATGCAATGCCCGTGGATGGCA 2334

Qy 794 SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer 813

Db 2335 TCCAAATTACTGCAAGCCCTAAGCTTCCAGCTTTCAGGGTCTGCTCTATCAATCTTTCA 2394

Qy 814 AspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArg 833

Db 2395 GATCGTCTGAACCAAGTGCTGAGGTGGCTTTAGTTTCAGTTGAGATTCTCTTCAGTCGG 2454

Qy 834 HisCysProLeuThrTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyr 853

Db 2455 CATTTGCTCTATATGTTATGTTTACAAATGGGAGGCTAAAAATTTCTTGAGAGGTTTGC 2514

Qy 854 IleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuPro 873

Db 2515 GTGAACACCAACATCTACCTCATCACCTCATCTCTCTCTCACTGATTGTACATTGCCA 2574

Qy 874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle 893

Db 2575 GCGGTTTGTCTCTTCCACCAACCAAGTTTATTTATTTCTCAGATTAGTAACATTGCAAGTATA 2634

Qy 894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSer 913

Db 2635 TGGTTTCTGTCTCTCTCTCTCTCCATTTTCGCCACGGGTATCTAGAAAATGAGGTGGAGT 2694

Qy 914 GlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSer 933

Db 2695 GCGTAGGCATAGACGAATGTGGAGAAACAGCAGAGTTTTTGGGTCATTGGTGGAGTATCC 2754

Qy 934 AlaHisLeuPheAlaValPheGlnGlyLeuLysValLeuAlaGlyIleAspThrAsn 953

Db 2755 GCTCATTTATTCGCTGTGTTTCAAGGTATCTCTCAAGTCTCTGCGGTATTGACACAAAC 2814

Qy 954 PheThrValThrSerLysAlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPhe 973

Db 2815 TTCACAGTTACCTCAAAAGCTTCAGATGACAGCAGAGACTTTTGTCTGAGCTCTACTTGTTC 2874

Qy 974 LysTrpThrThrLeuLeuLeuProThrThrThrIleLeuIleIleAsnMetValGlyVal 993

Db 2875 AAATGGACAACTCTGATTTCCCAACAGCGCTCATCTCATGTAAACTTAGTGGAGTT 2934

Qy 994 ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGly 1013

Db 2935 GTTCAGAGAGTCTCTATGCTATCAACAGTGCATCAACATCATGGGAGCCACTCTTTGCT 2994

Qy 1014 LysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMet 1033

Db 2995 AAGTTGTTCTTTGCTGCTTGTGGTGAITGTTTCACTGTGACCTTCTCCTCAAGGGTTGATG 3054

Qy 1034 GlyArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIle 1053

Db 3055 GGTGCAGAGAACCGGACTCTTACCATTGTTGGTCTGGTCTGTTCTCTTGGCTTCTATC 3114

Qy 1054 PheSerLeuLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle 1073

Db 3115 TTCTGTTGTTGCTGGTTAGATTGATCCCTTCTACTAGCCGAGTCACTGGCCCGGACATT 3174

Qy 1074 GlnThrCysGlyIleAsnCys 1080

Db 3175 CTGGAATGTGGAAATCAACTGT 3195

```

RESULT 15
AF027174      3682 bp  mRNA  linear  PLN 07-FEB-1998
LOCUS
DEFINITION
Arabidopsis thaliana cellulose synthase catalytic subunit (Ath-B)
mRNA, complete cds.
ACCESSION
AF027174
VERSION
AF027174.1  GI:2827142
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 3682)
Arioli,T., Peng,L., Betzner,A.S., Burn,J., Wittke,W., Herth,W.,
Camilleri,C., Hofte,H., Plazinski,J., Birch,R., Cork,A., Glover,J.,
Redmond,J. and Williamson,R.E.
Molecular analysis of cellulose biosynthesis in Arabidopsis
Science 279 (5351), 717-720 (1998)
AUTHORS
Arioli,T.
TITLE
Direct Submission
JOURNAL
Submitted (29-SEP-1997) Plant Science Centre, Australian National
University, Acton, Canberra, ACT 200, Australia
FEATURES
Location/Qualifiers
1..3682
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
1..3682
/gene="Ath-B"
256..3453
/gene="Ath-B"
/feature="RSW1-like"
/codon_start=1
/product="cellulose synthase catalytic subunit"
/protein_id="AAC39336.1"
/db_xref="GI:2827143"
/translation="MESGETAGKPMKNIIVPOTQICSDNVGKTVGDGRFVACDTCISF
PVCPCYERHGRDGNQPCQCKRYKLGKSPAIKPDKDEGLADEGTFVFNYPQKEK
ISERMLGHLTRKGEEGEPQYDKREVSHNHLPLRTSQDTSGEFSAPKRLSVSST
IAGKRLPYSDVNOGPNRRIIDPVGLGNVAKMVERVKGWKKOEKNTGPFVSTQAAASR
GGVIDADSTIDALLDEARQPLSRKVSIPSSRINPYRWIMLRVLVILFLHYR
ITNPVNAFALWLSVI CEIWFALSKWLDQPKFPNREYVLDRLALRYDREGPSQ
LAAVDIFVSTVDPLKEPLVLTANTVLSILAVDIPVDRKVSIVFDDGAWLSPESLAET
SEFARKWVFCCKYSIEPRAPWYFAAKIDYLKQVTSFVKDRAMKREYEEFKIRI
NALVSKALKCEEGWMDGTWPNGNTGDHPGMIQVFLGQNGLDAREGNELPLRVY
SREKRFQHHKKAGAMALVRSAVLNTPGPFILNLDGDHYINNSKALREAMCFLMDP
NLGKQVCYVQPFQDFGIDKNDRYANRNTVPFDINLRGLDGIQGVVYVGTGCVENRTA
LYGVEPPIKVKHKKSLKLCGGSEKKNKAKESKKKSGRHTDSTVPVFNLDIE
EGVEGAGDEDEKALLMSQMLEKRFQGSFVASTLMENGGVPPSATPENLLEKAIHV
ISCGVEDSKDMEIGWIGVSTEDILTGFQKHARGRSIYCMKPLPAFGSAPINLS
DRLNQVLWALGSLDIFSRHCPIWYNGNRLKFLERFAYVNTTITPTISPLMYCT
LLAVCLFTNQFTIPOISNIASITWFLSFLSFATGILEMRWSGVGIDSWRNEQFQWVI
GVSAHLFAVQGLIKLVAGIDNFTVTSKASDEGDFAEILYLPKMTLLIPPTLLI
VNLGVVAGVSYAINSGYQSMCPGLKLPFAFWIVLHYLPFLKGLMGRQNRTPTIIVV
WSVLLASTFSLNWRIDFTSRVTGFDILEGGINC"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3682
Score: 4507.00 Matches: 836
Percent Similarity: 87.54% Conservative: 96
Best local Similarity: 78.35% Mismatches: 107
Query Match: 78.00% Indels: 26
DB: 8 Gaps: 8
US-09-900-237a-30 (1-1080) x AF027174 (1-3682)
QY 19 CysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCys 38

```

```

Db 313 TGCAGATCTGTAGTGACAAATGTTGGCAAGACTGTTGTATGGAGATCGTTTGGCTGT 372
QY 39 AspValCysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThr 58
Db 373 GATATTTGTTTCATTCCTCCAGTGTTCGCGCCTTGTACGAGTATGAGAGGAAGATGGAAT 432
QY 59 GlnAlaCysLeuGlnCysLysThrLysTyrLysArgHisArgLysSerProAlaIleArg 78
Db 433 CAATCTTGTCCTCAGTGCANAAACCAAGATACAAGAGGCTCAAAGGTAGTCTGTATTCCT 492
QY 79 GlyGluGluGlyAspThrAspAlaAspAspGlySer---AspPheAsnTyrProAla 97
Db 493 GGTGATAAAGACGAGGATGGCTTAGCTGATGAAGTACTGTTCAGTTCACATACCCT--- 549
QY 98 SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsn 117
Db 550 -----CAGAAGGAGGAAATTTTCAGAGCGGATGCTTGTGTTGGCATCTTACT 594
QY 118 ThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSer 137
Db 595 CGTGGAGGAGGAGAAATGGGGAACCCCAAGTATGATAA----- 636
QY 138 LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln--- 156
Db 637 -----GAGGTCTCTCACAATCATCTCTCGTCTCAGCAGACAGACAGAT 681
QY 157 MetSerGlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsn 176
Db 682 ACTTCAGGAGAGTTCCTGCTGCCTCACCTGAACGCCTCTCTGATCTTCTACTCGCT 741
QY 177 IleSerArgArgAlaProPheProTyr---ValAsnHisSerProAsnProSerArgGlu 195
Db 742 GGGGGAAGCGCCTTCCTTATTCATCAGATGTCATCAATCAATCAACCAATAGAGGATTGTG 801
QY 196 PheSerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLys 215
Db 802 GATCTGTTGGACTCGGGAATGTAGCTTGGAAAGGAGAGAGTGTGCTGGTGGAAAATGAAG 861
QY 216 GlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg 235
Db 862 CAAGAGAAGAACTACTGGTCTGTC-----AGCAGCAGCGGCTGCTCTGAA---AGA 909
QY 236 AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp 255
Db 910 GGTGAGTAGATATTGATGCCAGCACAGATATCTCAGCAGATGAGGCTCTGCTGATGAC 969
QY 256 GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr 275
Db 970 GAAGCGAGCGAGCTCTGTCAAGGAAAGTTCATTCCTTCATCAGGATCAATCCTTAC 1029
QY 276 ArgMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThr 295
Db 1030 AGAATGGTATTATGCTCGCGCTTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1089
QY 296 AsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPhe 315
Db 1090 AACCAGTGCANATGCCTTTCCTATGGCTGGTCTCTGTGATATGATGAGATCTGGTTT 1149
QY 316 AlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyr 335
Db 1150 GCCTTATCTCGATTTTGGATCAGTTTCCCAAGTGGTTTCTCTGTAACCGTGAAACCTAC 1209
QY 336 LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnIleAlaVal 355
Db 1210 CTCGACAGCGTGTCTTTAAGATATGATCGTGAAGGTGAGCCATCATCAGTTAGTCTGTT 1269
QY 356 AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr 375
Db 1270 GACATTTTCGTGAGTACTGTTGACCCCTTGAGGAGGCCACCCCTTGTACAGCCACACA 1329
QY 376 ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp 395

```

[illegible]

Search completed: August 22, 2004, 20:43:20
Job time : 10363 secs

840	YGRGLKELFAVINTTIYPLTSLPLLIIYCIPLAICLLTCKFIPEISNFASIWFI	SLFI	899
901	SIFATGILEMRGSGVIDEWRNEQVWIGGISAHLFVAVPQGLLKVL	AGIDTNTFTVTSKA	960
900	SIFATGILEMRGSGVIDEWRNEQVWIGGISAHLFVAVPQGLLKVL	AGIDTNTFTVTSKA	959
961	NDEGDFAEILYMPKWTLLIPPTTILIIINMVGVVAGTSAVAINSGY	OSWG	1020
960	SDSDGDFAEILYMPKWTLLIPPTTILIIINLVGVVAGISVAINSGY	OSWG	1019

QY	QVHLHFF	LRGLMG	GRQNR	PTIIV	VMAVLLAS	IFSLWVR	VDPTFKLAGN	QICGNC	108
DB	1020	VIVHLYPFL	KGMLGR	QNRPTI	VVMAVLLAS	IFSLWVR	DDPTFKTRV	GDTRTCGNC	1079

RESULT 2
 AAY84108 standard; protein; 1077 AA.
 XX
 AC AAY84108;
 XX
 DT 03-JUL-2000 (first entry)
 XX
 XX Amino acid sequence of a maize cellulose synthase.
 XX
 KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
 KW transgenic plant; plant breeding marker.
 XX
 OS Zea mays.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 494
 FT /note= "encoded by MGG"
 FT
 XX WO2006092706-A2.
 XX
 PD 24-FEB-2000.
 XX
 PF 16-AUG-1999; 99WO-US018760.
 XX
 PR 17-AUG-1998; 98US-0096822P.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
 XX
 PI WPI; 2000-224343/19.
 DR N-PSDB; AAZ99494.
 DR
 XX
 PT New genes which encode maize cellulose synthase polypeptides in plants
 PT useful for modulating the expression of cellulose synthase in plants and
 PT to produce transgenic plants expressing the novel protein.
 XX
 PS Claim 15; Page 95-97; 119pp; English.
 XX
 CC The present sequence represents a maize cellulose synthase polypeptide.
 CC The cellulose synthase can be used for the improvement of stalk quality
 CC for improved stand or silage. It also provides an increased concentration
 CC of cellulose in the pericarp, hardening the kernel and improving its
 CC handling ability. The sequences are used to produce transgenic plants and
 CC seeds expressing the cellulose synthase. The polynucleotide is used for
 CC modulating, preferably increasing, the level of the synthase in a plant
 CC cell. The plants are preferably monocots. The polynucleotide is also used
 CC as a probe or primer in the detection quantitation or isolation of gene
 CC transcripts. The probes are useful in detecting deficiencies in the level
 CC of mRNA in screenings for desired transgenic plant, for detecting or
 CC mutations in the gene, for monitoring upregulation of expression or
 CC changes in enzyme activity in screening assays of compounds, for
 CC detection of any number of allelic variants of the gene, or for use as
 CC molecular markers in plant breeding programs. The isolated nucleic acids
 CC of the present invention can also be used for recombinant expression of
 CC their encoded polypeptides or for use as immunogens in the preparation

SQ	Sequence 1077 AA;		
Query Match 93.7%; Score 5415.5; DB 3; Length 1077;			
Best Local Similarity 92.3%; Pred. No. 0;			
Matches 997; Conservative 46; Mismatches 34; Indels 3; Gaps 2;			
QY	1	MDGDADALSKSRHGAGDVQCI	CADGLGTLTLDGVDFTACDVCRFPVCRPCYEHERKEGTQA 60
Db	1	MEGDADGVKSGRRGGVQCI	CGDGVGTAGDVFAACDVCGFPVCRPCYEYERKDGTOA 60
QY	61	CLOCKTKYKRRHGSFAIRG	EGDDTDADGSDFNYPASGTEDQKQKIADRMRSWRMVTGG 120
Db	61	CPQCKTKYKRRHGSFAIRG	EEGDDTDAD--SDFNYLASGNEQKQKIADRMRSWRMNVGG 118
QY	121	SGNVGHPKYDSGETGLSKY	DSGEIPIRGYVPSVTNSQMSGEIPGASPDHMMSPGTGNTSRR 180
Db	119	SGDVERPKYDSGETGLTKY	DSGEIPRGYIPSVTNSQISGEIPGASPDHMMSPGTGNTICKR 178
QY	181	APFPYVNHSPNPSREFSGS	IGNVAKRVDGWMKQDKGAIPTMTNGTSLAPSEGRAATDI 240
Db	179	APFPYVNHSPNPSREFSGS	IGNVAKRVDGWMKQDKGTIPTMTNGTSLAPSEGRGVGDI 238
QY	241	DASTEYNMEDALLNDETRO	PLSRKVPDIASSKINPRMVIIVLRLVLSIFLHYRLTNTPVN 300
Db	239	DASTDYNMEDALLNDETRO	PLSRKVPDISSRINPRMVIIVLRLVLSIFLHYRIITNPVN 298
QY	301	AYPLMLLSVICBIWFALSW	ILDOFPKWPPIINRETYLDRALRYDREGESQSLAANDIFVS 360
Db	299	AYPLMLLSVICBIWFALSW	ILDOFPKWPPIINRETYLDRALRYDREGESQSLAANDIFVS 358
QY	361	TVDPKEPPIVITANTVLS	ILAVDYPDKVSCYVSDDGASMLTFDALAETSEFARKWVPFV 420
Db	359	TVDPKKEPPIVITANTVLS	ILAVDYPDKVSCYVSDDGAAMLTFDALAETSEFARKWVPFV 418
QY	421	KYDIEPRAPERYFOCKIDY	LKQVQPSFVKDRAMKREYEEFKIRINALVSKALKVPTEE 480
Db	419	KYKINTEPRAPERYFOCKI	DYLKQVHPSFVKDRAMKREYEEFKRVYNGVLVAKAQKYPVE 478
QY	481	GWIMQDGTWPNGNTRDHP	GMIVFLHSGSGGLDTGNEPLRLVYVSREKRPFGFQHHKAG 540
Db	479	GWIMQDGTWPNGNTRDHP	GMIVFLHSGSGGLDTGNEPLRLVYVSREKRPFGFQHHKAG 538
QY	541	ANNALVRVSAVLINGQYML	NDCDCHYINNSKAVREAMCFMDPNLGPQVCYVQPPQRPFDG 600
Db	539	ANNALVRVSAVLINGQYML	NDCDCHYINNSKALREAMCFMDPNLGRSVCYVQPPQRPFDG 598
QY	601	IDRNDRYANRNTVFFDIN	LRGLDGIQGPVYVGTGCVFNRTALGYEPPIKAKKPGFLASL 660
Db	599	IDRNDRYANRNTVFFDIN	LRGLDGIQGPVYVGTGCVFNRTALGYEPPIKQKGGFGLUSL 658
QY	661	CGGKKASKSKRSRSDKK	SKNKHVDSSVPVFNLEDEIEEGVEGAGFDDDEKSVLMSQMSLEK 720
Db	659	CGGKKASKSKK--GSDKK	SKQXHVDSVPVFNLEDEIEEGVEGAGFDDDEKSVLMSQMSLEK 717
QY	721	RFGQSAFVASTLMYGGVPS	TPESSLLKEAHTVISCYGEDKSEWTEGTGWIYGSVTE 780
Db	718	RFGQSAFVASTLMYGGVPS	TPESSLLKEAHTVISCYGEDKTEWTEGTGWIYGSVTE 777
QY	781	ILTGFKMHARGWRSVYCM	PKRPAPKGSAPINLSDRLNQVLRWALGSVEILFSRHCPLMYG 840
Db	778	ILTGFKMHARGWRSYCM	PKRPAPKGSAPINLSDRLNQVLRWALGSVEILFSRHCPLMYG 837
QY	841	YGGRLKFLERFAYINTTII	YPLTSLPLLVYCIPLPAICLLTGKFIEMPEISNLASIWFIALFL 900
Db	838	YGGRLKFLERFAYINTTII	YPLTSLPLLVYCIPLPAICLLTGKFIETPEISNFIASIWFI 897
QY	901	SIFATGILEMRWSGVGID	EWNRNEQFVIGGISAHLPFAVFOGLLKVLAGIDTFTFTVTSKA 960
Db	898	SIFATGILEMRWSGVGID	EWNRNEQFVIGGISAHLPFAVFOGLLKVLAGIDTFTFTVTSKA 957
QY	961	NDEEGDFAELYMFKWTLLI	PPPTTILINMGVNVAGTSYAINSGYQSWGPLFGKLPFAFW 1020
Db	958	SDEEGDFAELYMFKWTLLI	PPPTTILINMGVNVAGISYAINSGYQSWGPLFGKLPFAFW 1017

Qy	1021	VIVHLYPFLKGLMGQRNRTPTIVVWAVLLASIFSLLVWRVDPFETRLAGPNITQCGINC	1080
Db	1018	VIVHLYPFLKGLMGQRNRTPTIVVWAVLLASIFSLLVWRIDPFTTRVTRVGPDTQCGINC	1077
RESULT 4			
ID	AAV84119	standard; protein; 1043 AA.	
XX	AAV84119;		
XX	03-JUL-2000	(first entry)	
XX	Amino acid sequence of a maize cellulose synthase.		
XX	Maize; cellulose synthase; stalk quality; stand; silage; cellulose;		
KW	transgenic plant; plant breeding marker.		
XX	Zea mays.		
XX			
EH	Key	Location/Qualifiers	
FT	Misc-difference 494	/note= "encoded by MCG"	
FT			
XX			
PN	W0200009706-A2.		
XX			
PD			
XX			
PF	16-AUG-1999;	99WO-US018760.	
PR	17-AUG-1998;	98US-0096822P.	
XX			
PA	(PION-) PIONEER HI-BRED INT INC.		
XX			
PI	Dhugga KS, Helentjaris TG, Bowen BA, Wang X;		
XX			
DR	WPI; 2000-224343/19.		
XX			
DR	N-PSDB; AAZ99527.		
XX			
PT	New genes which encode maize cellulose synthase polypeptides in plants		
PT	useful for modulating the expression of cellulose synthase in plants and		
PT	to produce transgenic plants expressing the novel protein.		
XX			
PS	Claim 15; Page 181-183; 119pp; English.		
XX			
CC	The present sequence represents a maize cellulose synthase polypeptide.		
CC	The cellulose synthase can be used for the improvement of stalk quality		
CC	for improved stand or silage. It also provides an increased concentration		
CC	of cellulose in the pericarp, hardening the kernel and improving its		
CC	handling ability. The sequences are used to produce transgenic plants and		
CC	seeds expressing the cellulose synthase. The polynucleotide is used for		
CC	modulating, preferably increasing, the level of the synthase in a plant		
CC	cell. The plants are preferably monocots. The polynucleotide is also used		
CC	as a probe or primer in the detection quantitation or isolation of gene		
CC	transcripts. The probes are useful in detecting deficiencies in the level		
CC	of mRNA in screenings for desired transgenic plant, for detecting or		
CC	mutations in the gene, for monitoring upregulation of expression or		
CC	changes in enzyme activity in screening assays of compounds, for		
CC	detection of any number of allelic variants of the gene, or for use as		
CC	molecular markers in plant breeding programs. The isolated nucleic acids		
CC	of the present invention can also be used for recombinant expression of		
CC	their encoded polypeptides or for use as immunogens in the preparation		
CC	and/or screening of antibodies. The proteins can be employed in assays		
CC	for enzyme agonists or antagonists of enzyme function or for use of		
CC	immunogens or antigens to obtain antibodies specifically immunoreactive		
CC	with a protein		
XX			
SQ	Sequence 1043 AA;		

Query Match 91.0%; Score 5255.5; DB 3; Length 1043;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 969; Conservative 42; Mismatches 32; Indels 3; Gaps 2;

Qy 1 MDGDADALKSGRHGAGDVQCICADGLGTTLDGDVFTACDVCRCFPVCPCEYHERKEGTQA 60
Db 1 MEGDADGKSGRRGGVQCICGDSGVTGTAAGDVFACDVCRCFPVCPCEYERKDGTA 60
Qy 61 CLOCKTKYKXHRGSPAIRGEGDDTDADGSDFNYPASGTEDQOKIADRMRSWRMTGG 120
Db 61 CPQCKTKYKXHRGSPAIRGEGDDTDAD--SDFNYLASGNEQOKIADRMRSWRMVGG 118
Qy 121 SGNVGHPKYDSGETGLSKYDSGEIPRGVSPVTSQMSGEIPGASPDHMMSPGTNISR 180
Db 119 SGDVGRPKYDSGETGLTKYDSGEIPRGVSPVTSQISGEIPGASPDHMMSPGTNIGKR 178
Qy 181 APPEYVNHSPNPSREFSGSICGNVAKKRVDCGKWKQDKGATPMTNGTSIAPSEGRAATI 240
Db 179 APPEYVNHSPNPSREFSGSICGNVAKKRVDCGKWKQDKGTIPMTNGTSIAPSEGRGVDI 238
Qy 241 DASTEYNMEDALLNDETRQPLSRKVPPLSSRINPYRMVIVLRLVLSIFLHYRLTNPVN 300
Db 239 DASTIDNEMEDALLNDETRQPLSRKVPPLSSRINPYRMVIVLRLVLSIFLHYRLTNPVN 298
Qy 301 AYPILWLLSVICEIWFALSWILDQPKPFPINRETYLDRALRYDREGEPSQLAADVIFVS 360
Db 299 AYPILWLLSVICEIWFALSWILDQPKPFPINRETYLDRALRYDREGEPSQLAADVIFVS 358
Qy 361 TVDPLKEPPIVTANTVLSILAADVDPVKVSCVSDGASMLTFDALAETSEFARKWVPFV 420
Db 359 TVDPMKEPPLVTANTVLSILAADVDPVKVSCVSDGGAMLTFDALAETSEFARKWVPFV 418
Qy 421 KKYDIEPRAPPEFYFCQKIDYLDKQVPSFVDRAMKREYEEFKIRINALYSKALKVPEE 480
Db 419 KKYNIERAPPEWYFSQKIDYLDKQVHPSFVDRAMKREYEEFKRVNGLVAKAQKPEE 478
Qy 481 GWINQDGTWPNNTRHPGMIQVFLGHSGLDTEGNEPLRLVYVREKRPFGOHKKAG 540
Db 479 GWINQDGTWPNNTRXDPHGMIQVFLGHSGLDTEGNEPLRLVYVREKRPFGOHKKAG 538
Qy 541 ANNALVRSVAVLTNGQVYMLNLDCHYTNNSKAVREACFLMDPNLGPQCVVQRPQPDG 600
Db 539 ANNALVRSVAVLTNGQVYMLNLDCHYTNNSKALREACFLMDPNLGRSVVQVQRPQPDG 598
Qy 601 IDRNDYANRNTVFDFINRLGLDGIQGPVYVGTGCVFNRTAIYGVPEPIKAKPGFLASL 660
Db 599 IDRNDYANRNTVFDFINRLGLDGIQGPVYVGTGCVFNRTAIYGVPEPIKAKPGFLSSL 658
Qy 661 CGGKKKASKKRSDDKKSKNKHVDSSVPVNLEDIEBGEVAGPDDPKSVLMQSMSLEK 720
Db 659 CGGKKKASKK-K-GSDKKSKQKHVDSSVPVNLEDIEBGEVAGPDDPKSVLMQSMSLEK 717
Qy 721 RFGQSAFVASTLMYEGVGPSSSTPESILLKEAIIHVISGVEDKSEWGTGFIWIGSVTDE 780
Db 718 RFGQSAFVASTLMYEGVGPSSSTPESILLKEAIIHVISGVEDKTEWGTGFIWIGSVTDE 777
Qy 781 ILTGFKKHARGWSVYCMKPEPAFKGSAPINLSRLNQLRWALGSEVILFSRHCPLWYG 840
Db 778 ILTGFKKHARGWSIYCMKPEPAFKGSAPINLSRLNQLRWALGSEVILFSRHCPLWYG 837
Qy 841 YGBRLKFLERPAYINTIYPTSLPLLVYCIILPAICLLTGKFIIMPETISNLASIWFIALL 900
Db 838 YGBRLKFLERPAYINTIYPTSLPLLVYCIILPAICLLTGKFIIMPETISNLASIWFIALL 897
Qy 901 SIFATGILEMWSVGVIDEWRNEQFVIGGISAHLEFVQGLLKLVLGADITNFTVTSKA 960
Db 898 SIFATGILEMWSVGVIDEWRNEQFVIGGISAHLEFVQGLLKLVLGADITNFTVTSKA 957
Qy 961 NDEEGDFAELNFKWTTLLIPIPTTILINIMVGVAGTSYAINSGVQSGPLFGKLFFAFW 1020
Db 958 SDEGDFAELNFKWTTLLIPIPTTILINILGVVAGISYAINSGVQSGPLFGKLFFAFW 1017
Qy 1021 VIVHLYPPLKGLMGQRNTPPTIVIV 1046
Db 1018 VIVHLYPPLKGLMGQRNTPPTIVIVW 1043

RESULT 5
AA184121
ID AAY84121 standard; protein; 1076 AA.
XX
AC AAY84121;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT Misc-difference 425
FT /note= "encoded by ANG"
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
PR 17-AUG-1998; 98US-0096822P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
DR WPI; 2000-224343/19.
DR N-PSDB; AAZ99533.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
XX to produce transgenic plants expressing the novel protein.
PS Claim 15; Page 196-199; 119pp; English.
XX
CC The present sequence represents a maize cellulose synthase polypeptide.
CC The cellulose synthase can be used for the improvement of stalk quality
CC for improved stand or silage. It also provides an increased concentration
CC of cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants and
CC seeds expressing the cellulose synthase. The polynucleotide is used for
CC modulating, preferably increasing, the level of the synthase in a plant
CC cell. The plants are preferably monocots. The polynucleotide is also used
CC as a probe or primer in the detection quantitation or isolation of gene
CC transcripts. The probes are useful in detecting deficiencies in the level
CC of mRNA in screenings for desired transgenic plant, for detecting or
CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
XX with a protein
SQ Sequence 1076 AA;

Query Match 89.0%; Score 5143; DB 3; Length 1076;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 947; Conservative 69; Mismatches 59; Indels 6; Gaps 6;
Qy 1 MDGDADALKSGRHGAGDVQCICADGLGTTLDGDVFTACDVCRCFPVCPCEYHERKEGTQA 60
Db 1 MDG-GDATNSGKHVAGVQCICGDSGVTGTAAGDGLFTACDVCRCFPVCPCEYERKDGTA 59


```
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.

PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 78.4%; Score 4530; DB 3; Length 1054;
Best Local Similarity 78.6%; Pred. No. 0;
Matches 839; Conservative 98; Mismatches 104; Indels 26; Gaps 8;

Qy 19 CQICADGLGTLTLDGVFTACDVCFVPCRCYEHKEGTACLOCKTKYKRRHGSFAIR 78
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 9 CQICSDNVGKTVDGDRFVACDICSFPVPCRCYERKDGNGQSCPCQCKTRYKRLKGSFAIP 68
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Qy 79 GEEGDDTDADGS-DENYPASGTEDQOKIADRMRSMNTGGSGNVGHPKYDSGEIGLS 137
| : : | | : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 69 GDXDEGLADEGTVFENYP-----QKEKISERMLGWHLTRGKEEMGEPOYDK----- 116
| : : | | : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Qy 138 KYDSGEIPRGYVPSVTSNQ-MSGHPIPGASPDHMMSPGNTISRAPPFY-VNHSNPNFSRE 195
| : : | | : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 117 -----EVSHNHLPLRTSRQDTSGEFSAASPERLSVSTIAGKRLPYSSDVNQSPNRRIV 171
| : : | | : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
```



```
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144325P.
PR 18-JUL-1999; 99US-0144352P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149317P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.

PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 78.4%; Score 4530; DB 3; Length 1065;
Best Local Similarity 78.6%; Pred. No. 0;
Matches 839; Conservative 98; Mismatches 104; Indels 26; Gaps 8;

Qy 19 CQICADGLTTLGDGVTACDVCRFPVCRPCYHERKEGTQACLOCKTKYKRHRGSPAIR 78
Db 18 CQICADGLTTLGDGVTACDVCRFPVCRPCYHERKEGTQACLOCKTKYKRHRGSPAIR 78
Qy 20 CQICADGLTTLGDGVTACDVCRFPVCRPCYHERKEGTQACLOCKTKYKRHRGSPAIR 79
Db 20 CQICADGLTTLGDGVTACDVCRFPVCRPCYHERKEGTQACLOCKTKYKRHRGSPAIR 79

Qy 79 GEGDDTDADDGS-DFNYPASGTEDQOKIADRMRSWRMNTGGSGNVGHPKYDSEIGLS 137
Db 80 GDXDEDGLADEGTVFNYP-----QKEKISERMLGHLTRGKEENGEPQYDK----- 127

Qy 138 KYDSGEIPRGVPSVTNSQ-MSGEIPGASPDHNMSTGNISRPAPPY-VNHSFNSRE 195
Db 128 -----EVSHNHLPRLTSRQDTSGEFSAASPERLSVSTIAGGKRLPYSSDVNQSPNRIV 182

Qy 196 FSGSIGNVAMKERVDMGKODKGAIPMTNGTSTAPSEGRAATDIDASTEXNMEDALLND 255
Db 183 DFGVLGNVAMKERVDMGKODKGAIPMTNGTSTAPSEGRAATDIDASTEXNMEDALLND 238

Qy 256 ETRQPLSRKVPFIASSKINPYRMVTLRLVLSIFLHYRLINPNVRNAYPLMLLSVCEIWF 315
Db 239 EARQPLSRKVPFIASSKINPYRMVTLRLVLSIFLHYRLINPNVRNAYPLMLLSVCEIWF 298

Qy 316 ALSWILDQFPKVPFIPINRETYLDRLALYDREGPSQLAAVDIFVSTVDPLKEPPIVTANT 375
Db 299 ALSWILDQFPKVPFIPINRETYLDRLALYDREGPSQLAAVDIFVSTVDPLKEPPIVTANT 358
```


Db 659 ESDKKSGRHTDSTVPVFNLDIEEGVEGAGFDDDEKALLMSQMSLEKRFQGSVAVFVASTL 718
 QY 734 MEYGVQSTPESLLKEAHVISCYEDKSEWTEIGWYGSVTEIDILGFKMHARGWR 793
 Db 719 MENGVPSPATPENLLKEAHVISCYEDKSDWGMGEIGWYGSVTEIDILGFKMHARGWR 778
 QY 794 SVYCMKRPAPKGSAPINLSDRNLQVLRWALGVSVEILFSRHCPWYGVGGRKFLERFAY 853
 Db 779 SIYCMKPLPAFKGSAPINLSDRNLQVLRWALGVSVEILFSRHCPWYGVGGRKFLERFAY 838
 QY 854 INTTIVLTSPLLVYICILPAICLLTGKTFIMPEISNLASIFIALFISFATGILEMRWS 913
 Db 839 VNTTIVPITSPLMYCTLPVCLFTNQFIIPQISNIASIFWLSLFSIFATGILEMRWS 898
 QY 914 GVGIDWNRNEQFVIGGSAHLFAVFGGLKVLVLAGIDTNTFTVTSKANDEEGDPAELMYF 973
 Db 899 GVGIDWNRNEQFVIGGSAHLFAVFGGLKVLVLAGIDTNTFTVTSKANDEEGDPAELMYF 958
 QY 974 KWTLLIPPTLLIINMVGVVAGTSYAINSGYQSGWPLFGKLFPAFWIVHLYPFLKGLM 1033
 Db 959 KWTLLIPPTLLIINMVGVVAGTSYAINSGYQSGWPLFGKLFPAFWIVHLYPFLKGLM 1018
 QY 1034 GRQNRTPVIVWVLLASIFSLWVRVDPPTTLRAGNIQTGCINC 1080
 Db 1019 GRQNRTPVIVWVLLASIFSLWVRVDPPTTLRAGNIQTGCINC 1065

RESULT 9

AAW33819
 ID AAW33819 standard; protein; 1065 AA.

AC AAW33819;

XX 17-OCT-2003 (revised)

DT 06-JUL-1998 (first entry)

XX Arabidopsis cellulose biosynthetic protein Ath-B.

XX Cellulose; cellulose synthase; RSW1 protein; beta-1,4-glucan;
 KW transgenic plant.

OS Arabidopsis thaliana; var. Columbia.

XX WO9800549-A1.

XX 08-JAN-1998.

XX 24-JUN-1997; 97WO-AU000402.

XX 27-JUN-1996; 96AU-00000699.

XX (AUSU) UNIV AUSTRALIAN NAT.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX Arioli A, Williamson RE, Betzner AS, Peng L;

XX WPI; 1998-086974/08.

XX N-PSDB; AAW06567.

XX DNA encoding cellulose biosynthetic enzyme - useful for manipulation of
 PT cellulose and beta-1,4-glucan.

XX Claim 29; Page 144-149; 207pp; English.

XX This polypeptide is encoded by cDNA clone Ath-B (see AAW06567) that was
 CC isolated from an Arabidopsis thaliana cDNA library using PCR primers (see
 CC AAT99632-34) based on cellulose synthase RSW1 genomic clone 23H12 (see
 CC AAW06563) and EST clone AAT20782 (see AAW06562). It is closely related to
 CC Arabidopsis cellulose synthase RSW1 protein (see AAW33816-17). Claimed
 CC nucleic acid molecules (see AAW06562-69) coding for claimed polypeptides
 CC (see AAW33816-20 and AAW46202) involved in cellulose biosynthesis can be
 CC used to manipulate the cellulose and/or beta-glucan content of transgenic
 CC plants. Expression of nucleic acids in the sense orientation increases

CC the level of cellulose and reduces the level of non-crystalline beta-1,4-
 CC glucan and starch, providing plants with modified strength and/or shape
 CC and/or fibre properties, or having increased resistance to stresses or
 CC pests. Antisense, ribozyme or co-suppression molecules can be used to
 CC reduce the cellulose content of a transgenic plant, e.g. to improve
 CC digestibility or to alter carbon partitioning such that increased carbon
 CC is available for growth, rather than deposited as cellulose. (Updated on
 CC 17-OCT-2003 to standardise OS field)

XX SQ Sequence 1065 AA;

Query Match 77.9%; Score 4499; DB 2; Length 1065;

Best Local Similarity 78.3%; Pred. No. 0;

Matches 835; Conservative 98; Mismatches 108; Indels 26; Gaps 8;

QY 19 CQICADGLGTLTGDDVFTACDVCPVPCPYEHERKEGTQACLOCKTKYKHRHGSFAIR 78
 Db 20 CQICSDNVGTVGDRFVACDICSFPVPCPYEYERDQNGQSCQCKTRYKRLKGSFAIP 79
 QY 79 GEEGDDTDADDGS--DENYPASGTEDOKKJADRMRSRMNTGGSGNVGHPKYDSGEICLS 137
 Db 80 GKDDEGLADEGTVEFNYP-----QKEKISERMLGWLHTRGKEEMGEPOYDK----- 127
 QY 138 KYDSGEIPRGYVPSVINSQ-MSGEIPGASPDHMHMSPTGMISRAPPFY-VNHSPNPSRE 195
 Db 128 -----EVSHNHLPLTSTRQDTSGEFSAPSPERLSVSSTIAGGKELPYSSDVNQSPNRV 182
 QY 196 FSGIGNVANKERVGDWGMKQDKGAIPTWNGTSTAPSEGRAATDIDASTEYVNMEDALIND 255
 Db 183 DPVGLGNVANKERVGDWGMKQDKNTGPV---STQAASE-RGVGDIDASTDILADEALIND 238
 QY 256 ETRQPLSRKYPASSKINPYRMVIVPLVLSIFLHYELTNPNVAYPLWLLSVICETWF 315
 Db 239 EAROLLSRKVSIPISSRINPYRMVIMLRVILCLFLHYRITNPVFNALMLVSVICETWF 298
 QY 316 ALSWILDQFPKWFPIINRETYLDRALRYDREGESQLAAVDIFVSTVDPLKEPPIVANT 375
 Db 299 ALSWILDQFPKWFVNRETYLDRLALRYDREGESQLAAVDIFVSTVDPLKEPPIVANT 358
 QY 376 VLSILAVDYPVKVSCYVSDDGASMLTFDALAETSEFARKWVPVKYKIDIEPRAPPEYFC 435
 Db 359 VLSILAVDYPVKVSCYVSDDGAAMLSFESLAETSEFARKWVPVKYKIDIEPRAPPEYFC 418
 QY 436 QKIDYLDKQVPSFKDRBRAMKYEYEFKIRINALVSKALKVPEEGIMODGTPKGNNT 495
 Db 419 AKIDYLDKQVTSFVKDRBRAMKYEYEFKIRINALVSKALKVPEEGIMODGTPKGNNT 478
 QY 496 RDHPGMIQVFLHSGGLDTEGNELPLRVYVSREKRPQFHHKKAGAMNALVRVSAVLTING 555
 Db 479 GDHEGMIQVFLGQNGGLDAEGNELPLRVYVSREKRPQFHHKKAGAMNALVRVSAVLTING 538
 QY 556 QYMLNDCDHYINNSKAVREAMCFMDPNLGPQVCYVQFPORFGIDRNDRYANRNTVFF 615
 Db 539 PFILNDCDHYINNSKALREAMCFMDPNLGLKQVCYVQFPORFGIDRNDRYANRNTVFF 598
 QY 616 DINLRGLDGIQGPVYVGTGCVFNRTALYGYEPPITKA--KKPGFLASLCGGKKKSKKR 673
 Db 599 DINLRGLDGIQGPVYVGTGCVFNRTALYGYEPPITKA--KKPGFLASLCGGKKKSKKR 658
 QY 674 SSDKKKSNKHVDSSVPVFNLEDEEGVEGAGFDDDEKALLMSQMSLEKRFQGSAAVFASTL 733
 Db 659 ESDKKSGRHTDSTVPVFNLDIEEGVEGAGFDDDEKALLMSQMSLEKRFQGSAAVFASTL 718
 QY 734 MEYGVQSTPESLLKEAHVISCYEDKSEWTEIGWYGSVTEIDILGFKMHARGWR 793
 Db 719 MENGVPSPATPENLLKEAHVISCYEDKSDWGMGEIGWYGSVTEIDILGFKMHARGWR 778
 QY 794 SVYCMKRPAPKGSAPINLSDRNLQVLRWALGVSVEILFSRHCPWYGVGGRKFLERFAY 853
 Db 779 SIYCMKPLPAFKGSAPINLSDRNLQVLRWALGVSVEILFSRHCPWYGVGGRKFLERFAY 838
 QY 854 INTTIVLTSPLLVYICILPAICLLTGKTFIMPEISNLASIFIALFISFATGILEMRWS 913

PR	05-AUG-1999;	99US-0147260P.	Best Local Similarity 80.4%; Pred. No. 0;	Matches 783; Conservative 80; Mismatches 92; Indels 19; Gaps 6;
PR	06-AUG-1999;	99US-0147303P.		
PR	06-AUG-1999;	99US-0147416P.		
PR	09-AUG-1999;	99US-0147493P.		
PR	09-AUG-1999;	99US-0147935P.		
PR	10-AUG-1999;	99US-0148171P.		
PR	11-AUG-1999;	99US-0148319P.		
PR	12-AUG-1999;	99US-0148341P.		
PR	13-AUG-1999;	99US-0148565P.		
PR	13-AUG-1999;	99US-0148684P.		
PR	16-AUG-1999;	99US-0149368P.		
PR	17-AUG-1999;	99US-0149175P.		
PR	18-AUG-1999;	99US-0149426P.		
PR	20-AUG-1999;	99US-0149722P.		
PR	20-AUG-1999;	99US-0149723P.		
PR	20-AUG-1999;	99US-0149929P.		
PR	23-AUG-1999;	99US-0149902P.		
PR	23-AUG-1999;	99US-0149930P.		
PR	25-AUG-1999;	99US-0150566P.		
PR	26-AUG-1999;	99US-0150884P.		
PR	27-AUG-1999;	99US-0151065P.		
PR	27-AUG-1999;	99US-0151066P.		
PR	30-AUG-1999;	99US-01511303P.		
PR	31-AUG-1999;	99US-0151438P.		
PR	01-SEP-1999;	99US-0151930P.		
PR	07-SEP-1999;	99US-0152363P.		
PR	10-SEP-1999;	99US-0153070P.		
PR	13-SEP-1999;	99US-0153758P.		
PR	15-SEP-1999;	99US-0154018P.		
PR	16-SEP-1999;	99US-0154039P.		
PR	20-SEP-1999;	99US-0154779P.		
PR	22-SEP-1999;	99US-0155139P.		
PR	23-SEP-1999;	99US-0155486P.		
PR	24-SEP-1999;	99US-0155659P.		
PR	28-SEP-1999;	99US-0156458P.		
PR	29-SEP-1999;	99US-0156596P.		
PR	04-OCT-1999;	99US-0157117P.		
PR	08-OCT-1999;	99US-0157753P.		
PR	08-OCT-1999;	99US-0157865P.		
PR	07-OCT-1999;	99US-0158029P.		
PR	08-OCT-1999;	99US-0158232P.		
PR	12-OCT-1999;	99US-0158369P.		
PR	13-OCT-1999;	99US-0159283P.		
PR	13-OCT-1999;	99US-0159294P.		
PR	13-OCT-1999;	99US-0159295P.		
PR	14-OCT-1999;	99US-0159329P.		
PR	14-OCT-1999;	99US-0159330P.		
PR	14-OCT-1999;	99US-0159331P.		
PR	14-OCT-1999;	99US-0159637P.		
PR	14-OCT-1999;	99US-0159638P.		
PR	18-OCT-1999;	99US-0159584P.		
PR	21-OCT-1999;	99US-0160741P.		
PR	21-OCT-1999;	99US-0160767P.		
PR	21-OCT-1999;	99US-0160768P.		
PR	21-OCT-1999;	99US-0160770P.		
PR	21-OCT-1999;	99US-0160814P.		
PR	21-OCT-1999;	99US-0160815P.		
PR	22-OCT-1999;	99US-0160980P.		
PR	22-OCT-1999;	99US-0160981P.		
PR	22-OCT-1999;	99US-0160989P.		
PR	25-OCT-1999;	99US-0161404P.		
PR	25-OCT-1999;	99US-0161405P.		
PR	25-OCT-1999;	99US-0161406P.		
PR	26-OCT-1999;	99US-0161359P.		
PR	26-OCT-1999;	99US-0161360P.		
PR	26-OCT-1999;	99US-0161361P.		
PR	28-OCT-1999;	99US-0161920P.		
PR	28-OCT-1999;	99US-0161992P.		
PR	28-OCT-1999;	99US-0161993P.		
PR	29-OCT-1999;	99US-0162142P.		
Query Match				
		72.9%;	Score 4209.5;	DB 3; Length 959;

DT 03-JUL-2000 (first entry)

XX Amino acid sequence of a maize cellulose synthase.

DE Maize; cellulose synthase; stalk quality; stand; silage; cellulose;

XX transgenic plant; plant breeding marker.

KW Zea mays.

OS

XX W0200009706-A2.

PN 24-FEB-2000.

XX

XX 16-AUG-1999; 99WO-US018760.

PF

XX 17-AUG-1998; 98US-0096822P.

PR

XX (PION-) PIONEER HI-BRED INT INC.

PA Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX WPI; 2000-224343/19.

XX N-PSDB; AAZ99506.

DR

XX New genes which encode maize cellulose synthase polypeptides in plants

PT useful for modulating the expression of cellulose synthase in plants and

PT to produce transgenic plants expressing the novel protein.

XX

PS Claim 15; Page 126-128; 119pp; English.

XX

CC The present sequence represents a maize cellulose synthase polypeptide.

CC The cellulose synthase can be used for the improvement of stalk quality

CC for improved stand or silage. It also provides an increased concentration

CC of cellulose in the pericarp, hardening the kernel and improving its

CC handling ability. The sequences are used to produce transgenic plants and

CC seeds expressing the cellulose synthase. The polynucleotide is used for

CC modulating, preferably increasing, the level of the synthase in a plant

CC cell. The plants are preferably monocots. The polynucleotide is also used

CC as a probe or primer in the detection quantitation or isolation of gene

CC transcripts. The probes are useful in detecting deficiencies in the level

CC of mRNA in screenings for desired transgenic plant, for detecting

CC mutations in the gene, for monitoring upregulation of expression or

CC changes in enzyme activity in screening assays of compounds, for

CC detection of any number of allelic variants of the gene, or for use as

CC molecular markers in plant breeding programs. The isolated nucleic acids

CC of the present invention can also be used for recombinant expression of

CC their encoded polypeptides or for use as immunogens in the preparation

CC and/or screening of antibodies. The proteins can be employed in assays

CC for enzyme agonists or antagonists of enzyme function or for use of

CC immunogens or antigens to obtain antibodies specifically immunoreactive

CC with a protein

XX

SQ Sequence 1074 AA;

Query Match 69.1%; Score 3993.5; DB 3; Length 1074;

Best Local Similarity 68.4%; Pred. No. 0;

Matches 749; Conservative 134; Mismatches 153; Indels 59; Gaps 20;

QY 2 DGDADA-LKSGRHGAGVQCICDGLGTLTLDGVFTACDVCRFPVCPYEHKEGTQA 60

DB 23 DGDAPVPAKPTKSANGVCQICGDTGVGSATGDFVACNECAFPVCPYEHKEGNQC 82

QY 61 CLOCTKYKRRGSPATRGEGDDTDADD-GSDFNYPASGTEDOKOKIADMRSRWNTG 119

DB 83 CPQCKTRYKRGSPRVHGDD-EEDVDLIDNEFN-----KQ----- 119

QY 120 GSGNVGHPKY----DSGEIGLS-KYDGEIPRGVPSVTN-SOMSGEIPGASPDHH-MMS 172

DB 120 --GNKGPEWQLQGDADLSSARHD----PHHRIPLTSCQISGIPDASPDHRSIR 173

QY 173 PTGN-ISRRAPFPYVNHSPNPSREF-SGSIGNVAWKRVDGWMKQDKGAIPMTNGTSIA 230

DB 174 PTSSYVDPSPVPV--RIVDPKDLNSVGLNSVDWKRVERSVKQDKNMLQVTN----- 226

QY 231 PSEGRATDIDASTEYNMEDALLNDETROPILSRKVPPIASSKINPYRMVIVILVLSIFL 290

DB 227 -KYPEARGDME-GTSGNGEDMQWVDARLPISRTVPISSNQLNRIYVILRLILCFFF 284

QY 291 HYRLTNVRNAYPLWLILSVICEWFWALSLLDOPPKWFFPINRETYLDRALRYDREGEPS 350

DB 285 QYRISHVRNAYGLNLVSVICEWFWALSLLDQPKWYPINRETYLDRALRYDREGEPS 344

QY 351 QLAADVIFVSTVDPLKEPPIVTANTVLSILAVDYPVDKVCYVSDDGASMLTFDALETS 410

DB 345 QLAPIDVFVSTVDPLKEPPLITANTVLSILAVDYPVDKVCYVSDDGASMLTPESLSETA 404

QY 411 EFARKWVPFVKKYVDIEPRAPEFFCOKIDYLDKQVQSFVKDRAMKREVEEFKIRINAL 470

DB 405 EFARKWVPFCKGNIEPRAPEFFFAQIDYLDKQIQSFVKERRAMKREVEEFKIRINAL 464

QY 471 VSKALVPEEGWTMODGTWPNGNTRDHPGMIQVFLGHSGLDTEGNELPRLVVSREKR 530

DB 465 VAKAQVPEEGWTMADGTAMPNPRDHPGMIQVFLGHSGLDTEGNELPRLVVSREKR 524

QY 531 PGFOHHKAGAMNALVRVSAVLINGQYMLNLDCHYINNKAVERAMCFMDPMLGPQVC 590

DB 525 PGFOHHKAGAMNALIRVSAVLINGAYLLNVDCDHYFNSSKALREAMCFMMDPALGRKTC 584

QY 591 YVOPPORFGIDRDRYANENTVFFDINRLDGIQGVVVGTCVFNRTAIYGEVEPIK 650

DB 585 YVOPPORFGIDILDRYANRNIVFFDINMGLDGIQGVVVGTCVFNRTAIYGEVEPIK 644

QY 651 AK--KPQFLASLGGKKASKKSKKSDKKSNKHVDSVPVFNLEIDIEEGVEGAGFDE 708

DB 645 EADLEPNIVKSCGRK-RKNKSYMDSQSRIMKRTESSAPIFNWEDIEEGIE--GYEDE 701

QY 709 KSVLMSQMSLEKRFQSGAAVASTLMEYGVGVSQSTPESLLKEAHIVISCGYEDKSEWT 768

DB 702 RSVLMSQRKLEKRFQSQPIFIASFMTQGGIPSTNPASLLKEAHIVISCGYEDKTEWGK 761

QY 769 EIGWIYGSVTEIITGFKMHARGWSVYCMKPAFKGSAPINLSDRNLQVLRWALGSVE 828

DB 762 EIGWIYGSVTEIITGFKMHARGWSVYCMKPAFKGSAPINLSDRNLQVLRWALGSVE 821

QY 829 ILSRHCPLWVGGRGLKFLERFAYINTTIVPLTSLPLLVVLCILPAICLTGKIMPEIS 888

DB 822 ILSRHCPIWYGYNGRLKLERLAYINTIVPITSVPLIAYCVLPALCLTNKEIIEIS 881

QY 889 NLSIMPIALFLSIPATGILEMRWSGVGIDEWARNQOFWVIGGISAHLFVAVFQGLLKVLA 948

DB 882 NYAGMFFILLFASIFATGILELRWSGVGIDEWARNQOFWVIGGISAHLFVAVFQGLLKVLA 941

QY 949 GIDTNFTVTSKANDEGDFAEIYMFKWTLLIPPTTILINMVCVAVAGTSVAINSGYSW 1008

DB 942 GIDTNFTVTSKANDEGDFAEIYVFKWTSLLIPPTTVLINLVMGVAIGSVAINSGYSW 1001

QY 1009 GPLFGKLFPAFWIVVHLYPELKGIMGRQNTPTTIVWAVLLASIFSLWVRVDPF--T 1065

DB 1002 GPLFGKLPFSIWIHLHLYPELKGIMGRQNTPTTIVWISILLASIFSLWVKIDPFISPT 1061

QY 1066 TRLAGPNIOTCGINC 1080

DB 1062 QKAAA--LGQCGVNC 1074

RESULT 13

AAAY84117

ID AAAY84117 standard; protein; 1074 AA.

XX

AC AAAY84117;

XX

DT 03-JUL-2000 (first entry)

XX Amino acid sequence of a maize cellulose synthase.

XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;

transgenic plant; plant breeding marker.

Zea mays.

WO200009706-A2.

24-FEB-2000.

16-AUG-1999; 99WO-US018760.

17-AUG-1998; 98US-0096822P.

(PION-) PIONEER HI-BRED INT INC.

Dhugga KS, Helentiaris TG, Bowen BA, Wang X:

WPI: 2000-224343/19.

New genes which encode maize cellulose synthase polypeptides in plants useful for modulating the expression of cellulose synthase in plants and to produce transgenic plants expressing the novel protein.

Claim 15: Page 165-167; 119pp; English.

The present sequence represents a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein

Sequence 1074 AA:

Query Match	69.1%;	Score 3993.5;	DB 3;	Length 1074;
Best Local Similarity	68.4%;	Pred. No. 0;		
Matches 749;	Conservative 134;	Mismatches 153;	Indels 59;	Gaps 20;

2 DGDADA-LKSGRHHGAGDVCQICADGLGTTLDGDVFTACDVCRFPVCRCPCYEHERKEGTQA 60

23 DGDAPVPAPKPTKSANGOVCOICGDTVGVSATGDFVACNECAFPVCRPCYEFYERKEGNOC 82

61 CLOCKTKYKRHRGSPAIRGEEGDDTDADD-GSDFNYPASGTEDQKQKIADRMRSWRMNTG 119

83 CPOCKTRYKROKSPRVHGDD-EEEDVDDLDNEFNY-----KO-----119

120 GSGNVGHPKY-----DSGEIGLS-KYDSGEIPRGVVPSTN-SOMSGEIPGASPDHH-MMS 172

```

120  --GNGKGPEWLOGDDADLSSARHD---PHHRIPRLTSGOOISGEIPDASPDRHSIRS 173

```

173 PTGN-ISRRAPFPVNHSPNPSREF-SGSIGNVAWKERVDPGWKMKODKGAIPMTNGTSTA 230

174 PTSSVDPSPVPV--RTVDPSPKDLNSYGI.NSVDPWKERVFSWRVKODKNMIQVTN----- 226

231 PSEGRAATDIDASTEYNMEDALI.NDETROPI.SRKVPPIASSKTNPYRMVTVI.RI.VVI.SIFL. 290

227 - KYPEARQDME-GTGSNGEDMOMVDDARI.PI.SRIVPTSSNOI.NI.VR.VI.II.PI.II.CFFF 284

FT	Misc-difference 245
FT	/note= "encoded by NAT"
XX	
PN	WO200009706-A2.
XX	
PD	24-FEB-2000.
XX	
PF	16-AUG-1999; 99WO-US018760.
XX	
PR	17-AUG-1998; 98US-0096822P.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	
PI	Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX	
DR	WPI; 2000-224343/19.
DR	N-PSDB; AAZ99530.
XX	
PT	New genes which encode maize cellulose synthase polypeptides in plants
PT	useful for modulating the expression of cellulose synthase in plants and
PT	to produce transgenic plants expressing the novel protein.
XX	
PS	Claim 15; Page 188-191; 119pp; English.
XX	
CC	The present sequence represents a maize cellulose synthase polypeptide.
CC	The cellulose synthase can be used for the improvement of stalk quality
CC	for improved stand or silage. It also provides an increased concentration
CC	of cellulose in the pericarp, hardening the kernel and improving its
CC	handling ability. The sequences are used to produce transgenic plants and
CC	seeds expressing the cellulose synthase. The polynucleotide is used for
CC	modulating, preferably increasing, the level of the synthase in a plant
CC	cell. The plants are preferably monocots. The polynucleotide is also used
CC	as a probe or primer in the detection quantitation or isolation of gene
CC	transcripts. The probes are useful in detecting deficiencies in the level
CC	of mRNA in screenings for desired transgenic plant, for detecting or
CC	mutations in the gene, for monitoring upregulation of expression or
CC	changes in enzyme activity in screening assays of compounds, for
CC	detection of any number of allelic variants of the gene, or for use as
CC	molecular markers in plant breeding programs. The isolated nucleic acids
CC	of the present invention can also be used for recombinant expression of
CC	their encoded polypeptides or for use as immunogens in the preparation
CC	and/or screening of antibodies. The proteins can be employed in assays
CC	for enzyme agonists or antagonists of enzyme function or for use of
CC	immunogens or antigens to obtain antibodies specifically immunoreactive
CC	with a protein
XX	
SQ	Sequence 1075 AA;
	Query Match 68.9%; Score 3981; DB 3; Length 1075;
	Best Local Similarity 68.5%; Pred.No.0;
	Matches 74%; Conservative 132; Mismatches 163; Indels 48; Gaps 19;
Qy	2 DGA-DALKSGRHGAGVCOICADGLGTLDGDVFTACDVCRCPCYHERKEGQA 60 : : : : : : : : : : : : : : : : : :
Db	23 DGDVPGSAKPTKSANGVCQICGSVGVSATGDFVACNECAFPVCRPCYERKEGNQC 82 : : : : : : : : : : : : : : : : : :
Qy	61 CLOCKTKYKRHRGSPAIRGEEDTDDAD-GSDFNYPASGTEDOKQKIADRMSRWMTG 119 : : : : : : : : : : : : : : : : : :
Db	83 CPQCKTRYKKQKGSPRVHGDE-DEEDVDLDFNFI-----KQ----- 119 : : : : : : : : : : : : : : : : : :
Qy	120 GSGNVGHKPYDSGBIGLSKYDSGBIIPRGVPSVTN-SQMSGIPEGASDPDH-MMSPTGN- 176 : : : : : : : : : : : : : : : : : :
Db	120 GSGKGPWEQLQGDDADLTSSARHE-PHRIPLRTSGQQISGEIPDASPDRHSIRSPTSY 178 : : : : : : : : : : : : : : : : : :
Qy	177 ISRRAPPPYVNHSNPSEEF-SGSIGNVAWKERVDCWKMKODKGAIPMNTNGTSTAPSEGR 235 : : : : : : : : : : : : : : : : : :
Db	179 VDPSPVPFVV--RIVDPSKDLNSYGINSVDWKERVESWRVKQDKNMQVTN----KYPEAR 232 : : : : : : : : : : : : : : : : : :
Qy	236 AATDIDASTEYNMEDALLNDETROPLSRKVPITASSKINPYRMWIVLRVLWLISFLHYELT 295 : : : : : : : : : : : : : : : : : :
Db	233 GG-DME-GTGSGENGEMQMWDARUPLSRIVPISSNQLNLRYVILRLIIICFFQFYR 290 : : : : : : : : : : : : : : : : : :
Qy	296 NEVRNAVPLWLLSVTCIEIWFALESWILDOPKPKWFPIINRETYILDRALRYDRGEPSQLAAV 355 : : : : : : : : : : : : : : : : : :

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 22, 2004, 17:19:28 ; Search time 172 Seconds

(without alignments)
3484.575 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MDGDADALKSRHGAGDVCO.....VDPFTTLAGNITCTGCNG 1080

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0900237/runat_18082004_081515_14658/app.query.fasta_1.1223
-DB-Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0900237 @cgn_1_84 @runat_18082004_081515_14658 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PTCUS.COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	4499	77.9	3614	4	US-09-221-013A-9
2	3963.5	68.6	3603	4	US-09-221-013A-5
3	3959.5	68.5	3673	4	US-09-221-013A-11
4	3633.5	62.9	3828	4	US-09-221-013A-7
5	3518.5	60.9	8411	4	US-09-221-013A-3
6	3477	60.2	3328	3	US-08-960-048-1
7	3477	60.2	3328	4	US-09-838-586-1
8	2629.5	45.5	2248	4	US-09-221-013A-1
9	2539.5	44.0	5009	4	US-09-221-013A-4
10	1833	31.7	4612	3	US-08-960-048-2
11	1833	31.7	4612	4	US-09-838-586-2
12	1756.5	30.4	1741	4	US-09-221-013A-13

13	478	8.3	281	4	US-09-313-294A-6485	Sequence 6485, Ap
14	477	8.3	300	4	US-09-313-294A-4753	Sequence 4753, Ap
15	442	7.6	291	4	US-09-313-294A-5519	Sequence 5519, Ap
16	401.5	6.9	268	4	US-09-313-294A-849	Sequence 849, App
17	389.5	6.7	286	4	US-09-313-294A-4696	Sequence 4696, Ap
18	349	6.0	284	4	US-09-313-294A-3993	Sequence 3993, Ap
19	312	5.4	287	4	US-09-313-294A-5274	Sequence 5274, Ap
20	283	4.9	283	4	US-09-313-294A-1981	Sequence 1981, Ap
21	279	4.8	278	4	US-09-313-294A-4015	Sequence 4015, Ap
22	265	4.6	317	4	US-09-313-294A-7123	Sequence 7123, Ap
23	251.5	4.4	2841	4	US-09-489-039A-3885	Sequence 3885, Ap
24	246	4.3	9540	1	US-07-689-008-1	Sequence 1, Appli
25	241	4.2	159	4	US-09-313-294A-1693	Sequence 1693, Ap
26	241	4.2	16836	4	US-08-147-236-1	Sequence 1, Appli
27	241	4.2	16836	4	US-09-147-236-10	Sequence 10, Appli
28	241	4.2	16836	4	US-09-522-474-1	Sequence 1, Appli
29	241	4.2	16836	4	US-09-522-474-10	Sequence 10, Appli
30	238.5	4.1	2262	1	US-07-689-008-3	Sequence 3, Appli
31	236.5	4.1	2217	4	US-09-543-681A-3356	Sequence 3356, Ap
32	230	4.0	2652	4	US-09-543-681A-2775	Sequence 2775, Ap
33	210.5	3.6	2112	4	US-09-489-039A-3649	Sequence 3649, Ap
34	209	3.6	2912	4	US-08-541-939-1	Sequence 1, Appli
35	209	3.6	2912	5	FCT-US91-01726-2	Sequence 2, Appli
36	195.5	3.4	1707	4	US-09-489-039A-3798	Sequence 3798, Ap
37	192.5	3.3	1863	4	US-09-489-039A-518	Sequence 518, App
38	189.5	3.3	250	4	US-09-313-294A-582	Sequence 582, App
39	136.5	2.4	2673	4	US-09-252-991A-2693	Sequence 2693, Ap
40	136.5	2.4	2715	4	US-09-252-991A-2686	Sequence 2686, Ap
41	136.5	2.4	2814	4	US-09-252-991A-2875	Sequence 2875, Ap
42	134	2.3	7430	4	US-08-956-171E-260	Sequence 260, App
43	133	2.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
44	131	2.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
45	128	2.2	2946	4	US-09-252-991A-7660	Sequence 7660, Ap

ALIGNMENTS

RESULT 1
US-09-221-013A-9
; Sequence 9, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3614
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (217)..(3411)
US-09-221-013A-9

Alignment Scores:
Pred. No.: 0
Score: 4499.00
Percent Similarity: 87.44%
Best Local Similarity: 78.26%
Query Match: 77.86%
DB: 4
Length: 3614
Matches: 835
Conservative: 98
Mismatches: 108
Indels: 26
Gaps: 8

US-09-900-237A-30 (1-1080) x US-09-221-013A-9 (1-3614)

QY 19 CysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCys 38
DB 274 TGCCAGATCTGTAGTGACAATGTTGGCAAGACTGTGTGATGGAGATCGTTTGTGCTGTGT 333
QY 39 AspValCysArgPheProValCysArgProCysTyrGluHisGluArgGlyGluGlyThr 58
DB 334 GATATTTGTTCAATCCAGTTGTGCGCCCTGCTAGAGTAGTAGAGGAAAGATGGGAAT 393
QY 59 GlnAlaCysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArg 78
DB 394 CAATCTGTCTCAGTGCAGAAACAGATACAGAGGCTCAAAGGTAGTCTCTGCTATTCT 453
QY 79 GlyGluGluGlyAspAspThrAspAlaAspAspGlySer---AspPheAsnTyrProAla 97
DB 454 GGTGATAAAGACGAGGATGGCTTGTAGTGAAGTACTGTGAGTTCAACTACCTCT--- 510
QY 98 SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsn 117
DB 511 -----CAGAAGGAGAAATTTTCAGAGCGGATGCTTGGTTGGCATCTTACT 555
QY 118 ThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSer 137
DB 556 CGTGGGAAGGAGAGGAAATGGGGAACCCAGTATGATAAA----- 597
QY 138 LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln--- 156
DB 598 -----GAGGTCTCTCACAAATCATCTTCTCTGCTCACGAGCAGACAAGAT 642
QY 157 MetSerGlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsn 176
DB 643 ACTTCAGGAGAGTTTCTGCTGCCACCTGAAAGGCTCTCTGATCTTCTACTATCGCT 702
QY 177 IleSerArgArgAlaProPheProTyr---ValAsnHisSerProAsnProSerArgGlu 195
DB 703 GGGGAAAGCGCTTCTCCCTATTTCATCAGATGTCAATCAATCACCAGAAATGAAGATGTG 762
QY 196 PheSerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLys 215
DB 763 GATCCTGTTGGACTCGGGAATGTAGCTTGGAAAGGAGAGTGTGATGGCTGGAAATGAAG 822
QY 216 GlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg 235
DB 823 CAAGAGAAGAATACTGGTCTCTGTC-----AGCACCGAGCTCTCTGAA---AGA 870
QY 236 AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp 255
DB 871 GGTGAGTAGATATTGATGCCAGCACAGATATCTTAGCAGATGAGGCTCTGCTGAATGAC 930
QY 256 GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr 275
DB 931 GAAGCGAGGACGCTCTGTGCAAGGAAGTTTCAATTCCTTCATCAGGATCAATCTCTTAC 990
QY 276 ArgMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThr 295
DB 991 AGAATGGTTATTATGCTCGGCTGTGTATCCCTTGTCTCTTGTGATTAACCTATAACA 1050
QY 296 AsnProValArgAsnAlaTyrProLeuThrLeuLeuSerValIleCysGlnIleTrpPhe 315
DB 1051 AAACCCAGTGCCTGCTCTGCTATGAGGCTGCTGCTGATGATGAGATCTGGTTT 1110
QY 316 AlaLeuSerTrpIleLeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyr 335
DB 1111 GCTTATCTCGAATTTGATCAGTTTCCCAAGTGGTTTCTGTGNAACCGTGNAACTTAC 1170
QY 336 LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaVal 355
DB 1171 CTCGACAGCTTGTCTTAAAGATATGATCGTGAAGGTGAGCCATCAAGTTAGTCTGTT 1230
QY 356 AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr 375
DB 1231 GACATTTCTGTAGTACTGTTGACCCCTTTGAAGGAGGCCACCCCTTGTGACGCCAACA 1290

QY 376 ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp 395
DB 1291 GTGCTCTCTATTCTGGCTGTGACTACCCAGTTGACAGGTGCTCTGTATTGTTCTGAT 1350
QY 396 AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys 415
DB 1351 GATGCTGCTGCTATGTTATCATTTGAATCCTTCAGAAACATCAGAGTTGCTCGTAAA 1410
QY 416 TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys 435
DB 1411 TGGGTACCATTTTGAAGAAATATAGCATAGACCTCGTGCACCAAGATGGTACTTGT 1470
QY 436 GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla 455
DB 1471 GCGAAAAATAGATTACTTGAAGGATTAAGTTGACACATCATTTGTCAAGATCGTAGAGT 1530
QY 456 MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu 475
DB 1531 ATGAAGAGGAATATGAGGAATTTAAATCCGAATCCAGATGCACTTGTTCCTCAAGCCCTA 1590
QY 476 LysValProGluGluGlyTrpIleMetGlnAspGlyThrProTyrProGlyAsnAsnThr 495
DB 1591 AAATGCTCTGAAGAGGCTGGGTTATGCAAGATGCGACACCGCTGGCAATTAATACA 1650
QY 496 ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu 515
DB 1651 GGGGACCATCCAGGAATGATCCAGGTCTCTTAGGGCAAAATGGTGGACTTCATGCAGAG 1710
QY 516 GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis 535
DB 1711 GCGAATGAGCTCCCGCTTGGTATATGTTCTCGAGAAAACGACAGGATTCAGCAC 1770
QY 536 HisLysLysAlaGlyValaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly 555
DB 1771 CACAAAAAGCTGGTCTATGAATGCCTGGTGAGAGTTTCAGCAGTTCTTACCAGTGA 1830
QY 556 GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu 575
DB 1831 CCTTTCATCTTGAATCTTGATTTGATCATTAATATAATAACAGCAAGCCCTTAAGAAA 1890
QY 576 AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro 595
DB 1891 GCAATGTGCTTCTGATGGACCAACCTCGGGAAGCAAGTTTGTATGTCAGTTCCCA 1950
QY 596 GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe 615
DB 1951 CAAAGATTGATGGTATGATAAGAACGATAGATATGCTAAATCGTAATACCGTGTCTTT 2010
QY 616 AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys 635
DB 2011 GATATTAACTTGAGAGTTTATAGTGGATTCAGGACCTGTATATGTCGAACTGGATCT 2070
QY 636 ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAla-----LysLys 653
DB 2071 GTTTTCAACAGAACAGCATTTATACGTTATGAACCTCCCAATAAAAAGTAAACACAGAAG 2130
QY 654 ProGlyPheLeuAlaSerLeuCysGlyGlyLysLysAlaSerLysSerLysLysArg 673
DB 2131 CCAAGTCTTTTATCAAGCTCTGTGTGATCAGAAAGAAAGAAATTCCAAGAGTAAAGAAA 2190
QY 674 SerSerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu 693
DB 2191 GAGTCGGAACAAAGAAATCAGGCAGGCATCTGACTCAACTGTTCTCTGATTCACCTC 2250
QY 694 GluAspIleGluGluGlyValGluGlyAlaGlyPheAspAspGluLysSerValLeuMet 713
DB 2251 GATGATAGAGAGGAGGATTAAGAGTGTGCTGTTTGTGATGATGAAAGCGCTCTTAATG 2310
QY 714 SerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeu 733
DB 2311 TCGCAATGAGCTCGAGAGCGATTTGGACAGTCTGCTGTTTGTGCTTCTTCTACCTTA 2370

```
QY 734 MetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIle 753
Db 2371 ATGGAAATGGTGGTGTCTCTCTCAGCAACATCCAGAAACATTTCTCAAGAGGCTATC 2430
QY 754 HisValIleSerCysGlyTyrGluAspLysSerGluTrpGlyThrGluIleGlyTyrIle 773
Db 2431 CATGTCTATTAGTTGTGTATGAGGATAAGTCAGATTGGGGAATGGAGATTGGATGCATC 2490
QY 774 TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArg 793
Db 2491 TATGGTTCTGTGACAGAAATATCTGACTGGGTTCAAAATGCGATGCCCGTGGATGCCGA 2550
QY 794 SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer 813
Db 2551 TCCATTACTGCGATGCCCTAAGCTTCCAGCTTTTCAAGGGTTCTGCTCTCATCAATCTTCA 2610
QY 814 AspArgLeuAsnGlnValLeuLeuArgTgtrAlaLeuGlySerValGluIleLeuPheSerArg 833
Db 2611 GATCGTCTGAACCAAGTGTGAGTGGGCTTTAGGTTTCAGTTGAGATTCTCTTCAGTCCG 2670
QY 834 HisCysProLeuTrpTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyr 853
Db 2671 CATGTCTTATGATGTTGTTCAATGGGAGGCTAAATTTCTTGAGAGGTTTGGCTAT 2730
QY 854 IleAsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuPro 873
Db 2731 GTGAACACCAACCATCTACCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2790
QY 874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluLysSerAsnLeuAlaSerIle 893
Db 2791 GCCGTTGTCTCTTCCACCAACGAGTTATTATTTCTCAGATTAGTAACATTCGCAAGTATA 2850
QY 894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSer 913
Db 2851 TGGTTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2910
QY 914 GlyValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSer 933
Db 2911 GCGGTAGGCAATACGAAATGGTGAGAAACGAGCAGTTTGGGTTCATTTGGTGGAGTATCC 2970
QY 934 AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn 953
Db 2971 GCTCATTTATTCGTGTGTTTCAAGGTATCTCTCAAGTCTCTTCCGCGTATTGACACAAC 3030
QY 954 PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe 973
Db 3031 TTCACAGTTTACCTCAAAAGCTTCAGATCAAGACGAGACTTTGCTGAGCTCTACTTCTTC 3090
QY 974 LysTrpThrThrLeuLeuLeuProProThrThrThrIleLeuIleAsnMetValGlyVal 993
Db 3091 AAATGGACAACACTTCTGATTCGCCCAACGACGCTGCTCATTTGTAACCTTAGTGGGAGTT 3150
QY 994 ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGly 1013
Db 3151 GTTGCAGGAGTCTCTATGATCTATCAACAGTGGATACCAATCATGGGACCACTCTTTGGT 3210
QY 1014 LysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMet 1033
Db 3211 AAGTTGTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3270
QY 1034 GlyArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeuAlaSerIle 1053
Db 3271 GGTGACAGAACCGGACTCTTACCATTTGTTGGTCTGGTCTGCTGCTGCTGCTGCTGCTGCT 3330
QY 1054 PheSerLeuLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle 1073
Db 3331 TTCTCGTTGTGGTTAGGATTGATTCCTTCTACTAGCCGAGTCACCTGGCCCGGACATT 3390
QY 1074 GlnThrCysGlyIleAsnCys 1080
Db 3391 CTGGAATGTGAATCAACTGT 3411
```

RESULT 2

```
US-09-221-013A-5
; Sequence 5, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 3603
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3243)
US-09-221-013A-5
Alignment Scores:
Pred. No.: 0 Length: 3603
Score: 3963.50 Matches: 733
Percent Similarity: 80.80% Conservative: 138
Best Local Similarity: 68.00% Mismatches: 168
Query Match: 68.60% Indels: 39
DB: 4 Gaps: 15
US-09-900-237A-30 (1-1080) x US-09-221-013A-5 (1-3603)
QY 2 AspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGluIle 21
Db 73 GATGGCGGGACCAACCTTTTGAAGATATGAAT-----GGCCAGATATGTCAGATC 123
QY 22 CysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCys 41
Db 124 TGTGGTGATGATGTTGACCTCGCTGAAACTGGAGATGCTTTGTCGCGGTAAATGATGT 183
QY 42 ArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGluAlaCys 61
Db 184 GCCTTCCCTGCTGCTCGGCTTGTATGATGATGACAGAGAAAGATGGAATCAGTGTTC 243
QY 62 LeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGluGlu 81
Db 244 CCTCAATGCAAGACTAGATTTCAGACACACAGGGGGAGTCTCTGTTGAGAGGATGAA 303
QY 82 GlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThrGlu 101
Db 304 GATGAGATGATGTTGATGATATCGAATGAGTTCATTAATACGCCAGGGAGCTAACAG 363
QY 102 AspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGlySer 121
Db 364 GCGAGACACCAACGCCCATGGCGAAGAGTTTCTTCT----- 399
QY 122 GlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp--- 140
Db 400 -----TCCTCTAGACATGAATTCACCAATTCCTCTTCTCACCCTATGGCCAT 447
QY 141 -----SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSer 158
Db 448 ACGTTTCTCGAGAGATTTCGACGCCCTGATACACAATCTGTGCGAACTACA-----TCA 501
QY 159 GlyGluIleProGlyAlaSerProAspHisHisMetSerProThrGlyAsnIleSer 179
Db 502 GGTCTCTTCTGACAGGAATGCTATTTCATCTCCA-----TATATTGAT 552
```

QY 179 ArgArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe---Ser 197
DB 553 CCACGGCAACCTGCTCCTGTA-----AGAACTCGTGGACCCGCTCAAAGACTTTGAACCTCT 606
QY 198 GlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAsp 217
DB 607 TATGGCTTGGTAAATCTTCACTGGAAAGAAAGAGTGAAGGCTGGAAGCTGAAGCAGGAG 666
QY 218 LysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAla 237
DB 667 AAAAATATGTTACAGATGACTCGT-----AAATACCATCAAGGGGAAGGAGGA 714
QY 238 ThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThr 257
DB 715 ---GAANTGAA--GGGACTGGTTCATGCGGAGAACTCCAATGGCTGATGATACA 768
QY 258 ArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMet 277
DB 769 GCTCTTCCTATGAGTCGTGTGTGCTATCCCATCTCTCGCCTAACCCCTTATCGGGTT 828
QY 278 ValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnPro 297
DB 829 GTGATATATCTCCGCTTATCACTGTGTGTTCTTTCGAAATATCGTACAACTCACCCCT 888
QY 298 ValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeu 317
DB 889 GTGAAAATATCATATCTTTGGTGTGACCTCGGTATCTGTGAGATCTGGTTTGCATTT 948
QY 318 SerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTrpLeuAsp 337
DB 949 TCTTGGCTTCTTGATCAGTTTCCCAATGGTACCCCATTTAAACAGGAGACTTATCTTGAC 1008
QY 338 ArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaValAspIle 357
DB 1009 COTCTCGCTTAAGATATGATCGACGCTGAACCATCACAGCTCGTCTCTTGTATGTTG 1068
QY 358 PheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeu 377
DB 1069 TTTGTGTAGTACAGTGGACCATTTGAAGAGCCCTCCCTGTTTACAGCAACACACAGTTCTC 1128
QY 378 SerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGly 397
DB 1129 TCGATTCCTTCTGTGGACTACCCGCTAGATAAAGTAGCCCTGTTATGTTTCAGATGATGT 1188
QY 398 AlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpVal 417
DB 1189 TCAGTATGCTTACCTTTGAATCCCTTCTGAAACCGCTGAGTTTGCAGAAATATGGGTA 1248
QY 418 ProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLys 437
DB 1249 CCAUTTTTCAAGAAATTCACACATTCAACCTAGGGCCCTGAAATCTATTTTGGCCAGAG 1308
QY 438 IleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLys 457
DB 1309 ATAGATTAATCTTGAAGCAAGATCAACCGCTCTTTTGTAAAGACGACGAGCTATGAAG 1368
QY 458 ArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysVal 477
DB 1369 AGAGAGTATGAAGAGTTTAAGTAGGATTAATGCTCTTGTTCGCAAGACACAGAAATC 1428
QY 478 ProGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAsp 497
DB 1429 CCTGAAGAGGCTGACCAATGACAGATGGTACTCCCTGGCCTGGTAAACACACATAGAGAT 1488
QY 498 HisProGlyMetIleGluValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsn 517
DB 1489 CATCTGGAATGATACAGGTGTTCTTAGCCCATAGTGGGGTCTGGATACCGATGGAAT 1548
QY 518 GluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLys 537
DB 1549 GAGTGCCTAGACATCATCTATGTTCTCGTGAAGAGCGCCTGGATTTCAACACACACAAA 1608
QY 538 LysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyr 557

DB 1609 AAGGCTGGAGCTATGAATGCAATGATCCGTGTATCTGCTGTTCTTACCAATGGAGCATAT 1668
QY 558 MetLeuAsnLeuAspCysAspHisTyrIleLeuAsnSerLysAlaValArgGluAlaMet 577
DB 1669 CTTTGTGAACGTGGATGTGATCATTTATTAACAGTAAGGCTATTAAAGAAGCTATG 1728
QY 578 CysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArg 597
DB 1729 TGTTCATGATGAGCCCGCTATTGGAAGAAAGTCTGTATGTCCAGTTCCCTCAACGT 1788
QY 598 PheAspGlyIleAspArgAspArgTyrValAlaAsnArgAsnThrValPhePheAspIle 617
DB 1789 TTTGACGCTATTGATTTGCACGATCGATATGCCAACAGAAATATAGTCTTTTTCGATAT 1848
QY 618 AsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPhe 637
DB 1849 AACATGAAGGGGTTGGATGATCCAGGGTCCAGTATATGTGGTACTGGTTGTGTTTT 1908
QY 638 AsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLys-----LysProGly 655
DB 1909 AATAGCAGGCTCTATATGGGTATGATCTGTTTTCAGCGAAGAGATTTAGAACCAAT 1968
QY 656 PheLeuAla---SerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSer 674
DB 1969 ATATTGTCAGAGAGCTGTTCGGGTCAAGGAAGAAAGGTAAAGGTAGCAAGAGATATAAC 2028
QY 675 SerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeuGlu 694
DB 2029 TACGAAAAGAGGAGGAGGCATCAACAGAGTGAATCTCAATGTCTCCACTTTTCAATATGGAG 2088
QY 695 AspIleGluGluGlyValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSer 714
DB 2089 GACATCGATGAGGTTTGGAA-----GGTTATGATGATGAGAGTCTATTCTAATGTC 2142
QY 715 GlnMetSerLeuGluLysArgPheGlyGlnSerAlaIlePheValAlaSerThrLeuMet 734
DB 2143 CAGAGGAGTGTAGAGAAGCGTTTGGTCAGTCGCGGTATTATTATGCGCAACCTTCATG 2202
QY 735 GluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHis 754
DB 2203 GAACAAGCGCGATCTCCACCAACCACTCCGCTACTCTTCTGAAGAGGAGCTATTCTAT 2262
QY 755 ValIleSerCysGlyTyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyr 774
DB 2263 GTTATAAGCTGTGGTTACAGACACAGACTGAATGGGCAAGAGATTTGGTTGATCTAT 2322
QY 775 GlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSer 794
DB 2323 GGTTCGCTGCGGAAGATATTCTTACTGGGTTCAAGATGTCATGCCCGGGTTGGATATCG 2382
QY 795 ValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAsp 814
DB 2383 ATCTACTCAATCTCTCCACGCCCTCGGTTCAAGGGATCTGCACCAATCAATCTTTCTGAT 2442
QY 815 ArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHis 834
DB 2443 CGTTTGAACCAAGTCTTCGATGGCTTTGGATCTATCGAGATTTCTTTAGCAGACAT 2502
QY 835 CysProLeuTrpTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIle 854
DB 2503 TGTCTCTATCTGATGGTTACCATGGAAGGTTGAGACTTTTGGAGAGATCGCTTATATC 2562
QY 855 AsnThrThrIleTyrProLeuProLeuSerLeuLeuValTyrCysIleLeuProAla 874
DB 2563 AACACATCGTCTATCTTATACATCCATCCCTCTTATTGGGTATGTATTCTTCCCGCT 2622
QY 875 IleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTrp 894
DB 2623 TTTTGTCTCATCCGACAGATTCATCATCCCGAGATAAGCAACTACGCGAGTATTTGG 2682
QY 895 PheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGly 914

```

Db      2683  TTCATTCTACTTCTCATCTCAATGCTGTGTGACTGGAATCTCTGGAGCTGAGATGAGCGGT 2742
Qy      915  ValGlyIleAspGluThrTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAla 934
Db      2743  GTGAGCATGTAGGATGTGGTGGAGAACGAGCAGTTCTGGGTCAATGTGGTGGACATCCGCC 2802
Qy      935  HisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPhe 954
Db      2803  CATCTTTTGTCTGCTTCCAAAGTCTACTTAAGTTCTTGTGATGATCGACCACTTC 2862
Qy      955  ThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLys 974
Db      2863  ACCGTTTACATCTAAAGCCACAGACGAAGATGGGATTTTGCAGAACTCTACATCTCAA 2922
Qy      975  TrpThrThrLeuLeuLeuProThrThrThrIleLeuLeuLeuLeuLeuMetValGlyVal 994
Db      2923  TGGACAGCTCTCTCAITCCACCAACACCGCTCTACTTGTGAACCTCATAGGCAATGTG 2982
Qy      995  AlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLys 1014
Db      2983  GCTGGTGTCTCTTAAGCTGTAACAGTGGCTTACCAGTCTGGTGGTCCGCTTTTCGGG 3042
Qy      1015  LeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGly 1034
Db      3043  CTCCTCTGGCTTATGGGTTATTGCCCATCTCTACCCCTTTCTTGAAAGTCTGTGGGA 3102
Qy      1035  ArgGlnAsnArgThrProThrIleValIleValIleValIleValIleValIlePhe 1054
Db      3103  AGACAAACCGAACCAACCAACATCGTCTGTTCTGTCTGTCTGTCTGTCTGTCTGTCT 3162
Qy      1055  SerLeuLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAsn 1072
Db      3163  TCGTTGCTTGGTTCAGATCAATCCCTTTGTGGAC-----GCCAATCCCAAT 3210

RESULT 3
US-09-221-013A-11
; Sequence 11, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3673
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (71)..(3313)
US-09-221-013A-11

Alignment Scores:
Pred. No.: 0 Length: 3673
Score: 3959.50 Matches: 732
Percent Similarity: 80.71% Conservative: 138
Best Local Similarity: 67.90% Mismatches: 169
Query Match: 68.53% Indels: 39
DB: 4 Gaps: 15

US-09-900-237A-30 (1-1080) x US-09-221-013A-11 (1-3673)

Qy      2  AspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIle 21

```

```

Db      143  GATGCGGGACCAACCTTTGAAGAATATGAAT-----GGCCAGATATGTCAGATC 193
Qy      22  CysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCys 41
Db      194  TGTGGTGATGATGTGGACTCGCTGAACTGGAGATGCTCTTTGTCGCGTGAATGAATGT 253
Qy      42  ArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAlaCys 61
Db      254  GCCTTCCTCGTGTCTGGCCTTGTATGAGTACGAGAGAAAGATGAACACTCAGTGTTC 313
Qy      62  LeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGluGlu 81
Db      314  CCTCAATGCAAGACTAGATTACAGACGACACAGGGGGAGTCTCTCGTGTGAAGAGATGA 373
Qy      82  GlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThrGlu 101
Db      374  GATGAGGATGATGTTGATGATATCGAGAATGAGTTCAATTACGCCAGGAGCTTAACAAG 433
Qy      102  AspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGlySer 121
Db      434  GCGAGACCAACGCCCATCGCGAAGAGTTTCTCT----- 469
Qy      122  GlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp--- 140
Db      470  -----TCCTCTAGACATGAATCTCAACCAATTCCTCTCTCACCATGGCCAT 517
Qy      141  -----SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSer 158
Db      518  ACGGTTTCTCGAGAGATTCGCACGCTGATACACAACTCTGCGAAGTCTTCA-----TCA 571
Qy      159  GlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSer 178
Db      572  GGTCTCTTG---GGTCTCTGACAGGAATGCTATTTCATCTCCA-----TATATTGAT 622
Qy      179  ArgArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe---Ser 197
Db      623  CCACGGCAACCTGTCCCTGTA-----AGAACTCGGACCGCTCAAAAGACTTGAACCTCT 676
Qy      198  GlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAsp 217
Db      677  TATGGGCTTGGTAATGTTGACTGGAAGAAAGAGTCTGAAGGCTGGAAGCTGAAGCAGGAG 736
Qy      218  LysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAla 237
Db      737  AAAAATATGTTACAGATGACTGGT-----AAATACCAGTGAAGGAAAGAGGAGA 784
Qy      238  ThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThr 257
Db      785  ---GAAATGGA---GGGACTGGTTCCTCAATGGCGAAGAACTCCAAATGGCTGATGATACA 838
Qy      258  ArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMet 277
Db      839  CGTCTCTATGATGCTGTGTGCTTATCCATCTCTCGCTTAAACCCCTTATCGGGTT 898
Qy      278  ValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnPro 297
Db      899  GTGATTATTCTCGGCTTATCATCTCTGTGTTTCTTGTCAATATCGTCAACCTCACCCCT 958
Qy      298  ValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeu 317
Db      959  GTGAAATAATGATATCTCTTGTGGTTCACCTCGGTTATCTGTGAGATCTGGTTGCATTT 1018
Qy      318  SerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAsp 337
Db      1019  TCTTGGCTTCTTGATCAGTTTCCAAATGGTACCCTTAACAGGAGACTTATCTTTGAC 1078
Qy      338  ArgLeuAlaLeuArgTyrAspArgGluGlyProSerGlnLeuAlaAlaValAspIle 357
Db      1079  CGTCTCGCTATAAGATATGATCGAGACGGTGAACCATCACAGCTCGTTCCTGTGATGTG 1138
Qy      358  PheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeu 377

```

Db 1139 TTTGTTAGTACAGTGGACCAATTTGAAAGAGCCCTCCCTTGTGTACAGCAAAACACAGTTCTC 1198
QY 378 SerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspGly 397
Db 1199 TCGATTCTTTCTGTGAGCTACCCGGTAGATAAGTAGCCTGTATGTATTTCAGATATGTT 1258
QY 398 AlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTyrVal 417
Db 1259 TCAGCTATGCTTACCTTTGAATCCCTTCTGAAACCGCTAGCTTTCGAAAGAAATGGGTA 1318
QY 418 ProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLys 437
Db 1319 CCATTTTGCAGAAATCAACATTCGAACCTAGGGCCCTGAAATCTTATTTTCCCAAGAG 1378
QY 438 IleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLys 457
Db 1379 ATAGATTACTTGAAGACCAAGATCCAACCGCTTTTGTGTAAAGAGCGAGCTATGAAG 1438
QY 458 ArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysVal 477
Db 1439 ACAGAGTATGAAGAGTTTAAAGTAGGATAAATGCTCTTGTGTCACAAAGCACAGAAATC 1498
QY 478 ProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGluAsnAsnThrArgAsp 497
Db 1499 CCTGAAGAGGCTGGACAAATGAGATGGTACTCCCTGGCCCTGGTAAACAACACTAGAGAT 1558
QY 498 HisProGlyMetIleGlnValPheLeuGlyHisSerGlyLeuAspThrGluGlyAsn 517
Db 1559 CATCTCGAATGATACAGGTGTTCTTAGGCCATAGTGGGGGTCTGGATACCGATGGAAAT 1618
QY 518 GluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLys 537
Db 1619 GAGCTGCCTAGACTCATCTATCTTCTGTGAAAGAGCGCCCTGGATTTCAACACACACAAA 1678
QY 538 LysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyr 557
Db 1679 AAGGCTGGAGCTATGAATGCATGTGATCCGTGTATCTGTGTGTTCTTACCAATGGACATAT 1738
QY 558 MetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMet 577
Db 1739 CTTTGTGAAGCTGATGTGATCATTTACTTTAATACAGTAAAGCTATTAAAGAGCTATG 1798
QY 578 CysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArg 597
Db 1799 TGTTTTCATGATGACCCGCTATTGGAAGAAGTGTCTATGTCTCAGTTCCTCAACGT 1858
QY 598 PheAspGlyIleAspArgAsnAspTyrAlaAsnArgAsnThrValPhePheAspIle 617
Db 1859 TTTGACGGTATTGATTGACCGATCGATATGCCAAGAGATATAGTCTTTTTCGATATT 1918
QY 618 AsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPhe 637
Db 1919 AACATGAAGGGTGGATGGTATCCAGGTCAGTATATGTGGTACTGGTGTGTGTTT 1978
QY 638 AsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLys-----LysProGly 655
Db 1979 AATAGCAGGCTCTATATGGGTATGATCTCTGTTTTCACGGAAGAAGATTTAGAACCAAT 2038
QY 656 PheLeuAla---SerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSer 674
Db 2039 ATATTGTCAAGAGCTGTTCGGGTCAAGGAAGAAGGTAAGTAGCAAGAAAGTATAC 2098
QY 675 SerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeuGlu 694
Db 2099 TACGAAAGAGGAGGAGGCATCAACAGAGTGACTCCAATGCTCCACTTTTCAATATGGAG 2158
QY 695 AspIleGluGluGlyValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSer 714
Db 2159 GACATCGATGAGGGTTTGA-----GGTTATGATGATGAGAGGTCTATTCTAATGTC 2212
QY 715 GlnMetSerLeuLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMet 734
Db 2213 CAGAGGAGTGTAGAGAGGTTTTGGTCACTGCGGTATTTATTTGCGGCAACCTTCATG 2272

QY 735 GluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHis 754
Db 2273 GAAACAGGCGGATTCACCAACCAACCAATCCCGTACTCTTCTGAAGAGGCTATTCTAT 2332
QY 755 ValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyr 774
Db 2333 GTTATAAGCTGTGGTTACGAACAAGACTGAATGGGGCAAGAGATCGTTGGATCTAT 2392
QY 775 GlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSer 794
Db 2393 GGTTCGGTACGGAAGATATCTTACTGGGTCAAGATGCATGCCGGGTGGATATCG 2452
QY 795 ValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAsp 814
Db 2453 ATCTACTGCAATCCCTCCAGCCCTGGTTCAGGGATCTGCACCAATCAATCTTCTGAT 2512
QY 815 ArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHis 834
Db 2513 CGTTTGAACCAAGTTCTTCGATGGCTTTGGGATCTATCGAGATTCTTCTTAGCAGACAT 2572
QY 835 CysProLeuTyrGlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIle 854
Db 2573 TGTCTTATCTGGTATGGTTACCATGGAAGTTGAGACTTTTGGAGAGGATCGCTTATATC 2632
QY 855 AsnThrThrIleTyrProLeuThrSerLeuLeuValTyrCysIleLeuProAla 874
Db 2633 AACACCATCGTCTATCTTATACATCCATCCCTCTTATTCGCTATTGATTTCTCCCGT 2692
QY 875 IleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyr 894
Db 2693 TTTTGTCTCATCACCAGACAGATTATATACCCGAGATAAGCAACTACGCGAGTATTTGG 2752
QY 895 PheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGly 914
Db 2753 TTCTTCTACTCTTCTCATCTCAATTGCTGTGACTTGGAACTCTGGAGCTGAGATGGAGCGT 2812
QY 915 ValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIleGlyIleSerAla 934
Db 2813 GTGACATTTGAGATTGGTGGAGGAACGAGCAGATTCTGGTTCATTTGGTGCAATCCGCC 2872
QY 935 HisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPhe 954
Db 2873 CATCTTTTGTCTGTCTCCAGGTCTACTTAAGGTTCTTGTCTGGTATCGACACCACTTC 2932
QY 955 ThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLys 974
Db 2933 ACCGTTTACATCTAAAGCCACAGAGAAATGGGAATTTTGCAGAACTCTACATCTTCAA 2992
QY 975 TrpThrThrLeuLeuIleProProThrThrIleLeuIleIleAsnMetValGlyValVal 994
Db 2993 TGGACAGCTCTTCTCATTTCCACCAACCCGCTTACTTGTGAACCTCATAGCATTTGG 3052
QY 995 AlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLys 1014
Db 3053 GCTGTGTCTCTTATGCTGTAAACAGTGGCTACCAAGTCTGGTGGGTCCGCTTTTCGGAAG 3112
QY 1015 LeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGly 1034
Db 3113 CTCTCTTCGCTTATGGGTATTGCCCATCTTACCCCTTTCTTGAAGGCTGTGTGGGA 3172
QY 1035 ArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePhe 1054
Db 3173 AGACAAACCGAAACCAACCAATCGTCTTGTGTCTGTCTTCTTCTCGCTCCATCTTC 3232
QY 1055 SerLeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGlyProAsn 1072
Db 3233 TCGTTGCTTTGGGTGAGGATCAATCCCTTTGTGGAC-----GCCAATCCCAAT 3280

RESULT 4

US-09-221-013A-7

; Sequence 7, Application US/09221013A

; Patent No. 6495740


```
QY 547 ArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyr 566
Db 1880 CGAGTCTCTGCTGTTCTATCAACCGCTCTTACCTTCTTAATGTCGATGTCATCACTAC 1939
QY 567 IleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGly 586
Db 1940 ATCAACAACAGCAAGCAATAGAGAACTATGTGTTCATGATGACCCGCAATCGGGA 1999
QY 587 ProGlnValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArg 606
Db 2000 AAGAAAGTTTGTATGTTTCAGTTCGCGAGAGATTGATGGGATTCATAGATAGATAGA 2059
QY 607 TyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGlyIleGln 626
Db 2060 TACTCAACCGTAACGTTGTGTTCTTGTATATTAACATGAAGGCTTCATGGGATACNA 2119
QY 627 GlyProValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGlu 646
Db 2120 GGACCGATATATGTCGGGACAGGTTGTGTGTTTAGAAAAACAGGCTCTTTATGCTTTGAT 2179
QY 647 ProProIleLysAlaLysLysProGlyPheLeuAlaSer----- 659
Db 2180 GCACCAAGAGAGAGAACCCACGAGCAAAACCTGTAACTGTGGCTAAATGGTGTGT 2239
QY 660 LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLys 679
Db 2240 TTGTGTTGGGTTGAGAAAG-----AAGAGTAAACGAAAGCCACAGATAAGAAACT 2293
QY 680 SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGly 699
Db 2294 AACACTAAA-----GAGACTTCAAGACAGATTTCATGCGCTAGAGAAATGTCGACGAAGT 2347
QY 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db 2348 GTTATCGTCCCGAGTGTCAAAATGTTGAGAGAGATCTGAAGCAACACAAATTTGAAATGGAG 2407
QY 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyVal 739
Db 2408 AAGAAGTTTGGACAATCTCCGGTTTCGTTGCTCTGCTGTCTACAGAACGGTGGAGTT 2467
QY 740 ProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGly 759
Db 2468 CCCCCTAAGCAAGCCCGCATGTTGTTTGAAGAGAGCCATTCAAGTTATTATAGCTGCGGG 2527
QY 760 TyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGlu 779
Db 2528 TACCAAGATAAAACCAATGGGAAAGAGATCGGGTGATTTATGGATCGGTGACTGAA 2587
QY 780 AspIleLeuThrGlyPheLysMethHisAlaArgGlyTyrArgSerValTyrCysMetPro 799
Db 2588 GATATCCTGACGGTTTCAAGATGCATTCGCCATGGATGGAGATCTGTGTACTGTATGCT 2647
QY 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
Db 2648 AAGCGTGAGCTTTTAAAGGATCTGCTCTCTATTAACTTGTACAGATCGTCTTCATCAAGTT 2707
QY 820 LeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyr 839
Db 2708 CTACGTTGGGCTCTTGGCTCTGTAGATTTCTTGACGACAGATTCGCGATATGGTAT 2767
QY 840 GlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
Db 2768 GGTATGTTGGTGGTTTAAATGGTTGGAGAGATTCTCTTACATCAACTCTGCTGCTAT 2827
QY 860 ProLeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
Db 2828 CCTTCGACTTCACTTCCATGTGATGCTATTGTTCTCCCGCGGTGTTGTTTACTCACA 2887
QY 880 GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPhe 899
Db 2888 GGAAATTAATTCGTCCTCGATAGATAGCAACTACGACAGGTATATCTCTTATGCTCATGTC 2947
QY 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGlu 919
```

```
Db 2948 ATATCATAGCAGTAACCTGGAATCCTCGAAATGCAATGGGAGGTGTCGGAATCGATGAT 3007
QY 920 TrpTrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaVal 939
Db 3008 TGGTGGAGAAACGAGCAGTTTTTGGGTAATCGGAGGGGCTCTCGCATCTATTGCTCTG 3067
QY 940 PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 959
Db 3068 TTTCAAGGTTTGTCTCAAAAGTTCTAGCCGAGGTAAACAGAAATTCACAGTCACTTCAAAA 3127
QY 960 AlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeu 979
Db 3128 GCACGAGAC---GATGGAGCTTCTCTGAGCTTTACATCTTCAAGTGGACAACCTTTGTTG 3184
QY 980 IleProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyr 999
Db 3185 APTCTCTCGACACACACTTCTGATCATTAAACATCATTTGGAGTTATTGTCGGCGTTTCTGAT 3244
QY 1000 AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3245 GCATTAGCAATGGCTATGACTCATGGGGACCTCTCTTTGGAGACTTTTCTTCGCTCTT 3304
QY 1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
Db 3305 TGGGTCAITGTTTCAITTTATACCATTCCTCAAGGGAATGCTTGGGAAGCAACAAAATG 3364
QY 1040 ProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1059
Db 3365 CCTACGATTTTGTGCTGTGCTATTCTTCTAGCTTCGATCTTGCACACTCTTTGTTGGTTC 3424
QY 1060 ArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
Db 3425 AGAATTAACCCGTTTGTGCTAAA---GGGGACCCAGTGTGGAGATCTGTGCTGAT 3481
QY 1080 Cys 1080
Db 3482 TGT 3484

RESULT 5
US-09-221-013A-3
; Sequence 3, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 8411
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-221-013A-3

Alignment Scores:
Pred. No.: 0 Length: 8411
Score: 3518.50 Matches: 743
Percent Similarity: 55.50% Conservative: 140
Best Local Similarity: 46.70% Mismatches: 158
Query Match: 60.83% Indels: 553
DB: 4 Gaps: 28

US-09-900-237A-30 (1-1080) x US-09-221-013A-3 (1-8411)
```

Qy	225	-----AsnGlyThrSerIle-	-----	229
Db	3965	AAGGAGGAGAAATTGAAGGGACTGGTTCCAAATGCGGAAGAAATCC	AAATGCTAAAGTGGAAA	4024
Qy	230	-----AlaProSerGlu-	-----GlyArgAlaAlaThrAspI	240
Db	4025	TACTAGACCAATATCTTTATTTGTCCTCAACTCAACACAGCTCTTGGCCGCTGATGCTAAATACC	-----	4084
Qy	240	leAspAlaSerThrGluThrAsnMetGluAspAlaLeu-	-----	252
Db	4085	ACTCTTG-GTTTCTTATTATGATTGATAGACATAATTAAGTATCTGCTTTGTTACATTT	4143	
Qy	253	-----LeuAsn-	-----AspGluThrArgGlnProL	261
Db	4144	GTTTCTCTCCACTCAATTATGTTCTCGTACTTACAGGGCTGATGATACACGCTCTTCCCTA	4203	
Qy	261	euSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleValL	281	
Db	4204	TGAGTCGTGGTGGCCCTATCCCACTTCTCGCCTAACCCCTTATCGGGTGTGATATTC	4263	
Qy	281	euArgLeuValValLeuSerIlePheLeuHi sTyrArgLeuThrAsnProValArgAsnA	301	
Db	4264	TCCGGCTTATCACTCTGTTTCTCTCTTCAATATCGTACAACTCACCCCTGTGAAAAATG	4323	
Qy	301	laTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIleL	321	
Db	4324	CATATCCTTTGTGGTTGACCTCGGTTATCTGTGAGATCTGGTTTGCATTTCTTGGCTTC	4383	
Qy	321	euAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAlaL	341	
Db	4384	TTGATCAGTTTCCCAAAATGGTACCCCAATTAAACAGGAGACTTATCTTTGACCGCTCTCGCTA	4443	
Qy	341	eu-	-----	341
Db	4444	T-AGGTTGGTCTTTAAGTTTATACATCCCTACTCTCATCTCTCTTTTATGTTAACT	4502	
Qy	341	-----	-----	341
Db	4503	TGATATCTCTATCACAGTTTTCGATAGTTGACTTTTTTCCCTCGTAAATTTAAATTTAAA	4562	
Qy	341	-----	-----	341
Db	4563	TTTAGACAATGGTGCATCTGAATTTTGATTATGATATATCTTAAGAAGATTATGATGTA	4622	
Qy	341	-----	-----	341
Db	4623	AATCTGAATTTAGTAGAAAAACCATCTGCCAATCTACTGACCATTGAAAGAGGCTC	4682	
Qy	341	-----	-----	341
Db	4683	GACTATGATAGAAGCATGCCAAGTGGAGTGTTTATTAAAGATAGAGCTTAGCTATTATACT	4742	
Qy	342	-----ArgTyrAspArgGluGlyGluP	349	
Db	4743	GAITTTATATGTGTTTGTGTTTCTTATTGTAGATATGATCGAGACGGTGAAC	4802	
Qy	349	roSerGlnLeuAlaAlaValAspIlePheValSerThrValAspProLeuLysGluProp	369	
Db	4803	CATCACAGCTCGTCTCTGTTGATGTGTTTGTGTAGTACGTGGACCCATTGAAGAGGCTC	4862	
Qy	369	rolIeValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysV	389	
Db	4863	CCCTTGTATACAGCAAAACACAGTCTCTCGATTCTTCTGTGGACTACCCGGTAGATAAAG	4922	
Qy	389	alSerCysTyrValSerAspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluTr	409	
Db	4923	TAGCCTGTATGTTTCAGATGATCGTTTCAGCTATGCTTACCTTTGAATCCCTTTCTGAAA	4982	
Qy	409	hrSerGluPheAlaArgLysTrpValProPheValLysTyrAspIleGluProArgA	429	
Db	4983	CCGCTGAGTTTGCAGAGAAATGGGTACCATTTTGCAGAAATTTCAACATTTGAACCTAGGG	5042	
Qy	429	laProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGlnProSerP	449	

Db 5043 CCCCCTGAATCTTATTTGTCAGAGAGATAGATTACTTGAAGCACAAGATCCACCGTCTT 5102
QY 449 heValLysAspArgAlaMetLys----- 457
Db 5103 TTGTTAAAGAGCGAGAGCTATGAAGGTCAATTGAAAAAGTCCACCTCTCTCATCCATA 5162
QY 458 -----ArgLysTyrGluGluP 463
Db 5163 CGGCAAGAGATTGACTGACTTTTTCTTTGGTTTGTATTGACAGAGAGAGTATGAAGGT 5222
QY 463 heLysileArgileAsnAlaLeuValSerLysAlaLeuLysValProGluGluGlyTyrPI 483
Db 5223 TTAAAGTGAGGATAAATGCTCTTGTGCCAAGCAGAGAAATCCCTGAAGAGGCTGGA 5282
QY 483 leMetGlnAspGlyThrProThrProGlyAsnAsnThrArgAspHisProGlyMetIle- 502
Db 5283 CAATGAGGATGGTACTCCCTGGCCTGGTAAACAACACTAGAGATCATCCTGGAATGATAC 5342
QY 502 ----- 502
Db 5343 AGGTACAGTGGCAATCCCTTGATTGTGACAGAGGATACGTAAGGAACAATGTTT 5402
QY 503 -----GlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluG 516
Db 5403 ACATCGTTTTGTTTCAAATTCAGGTGTTCTTAGGCCATAGTGGGGTCTGGATACCGATG 5462
QY 516 lysAsnLeuProArgLeuValTyrValSerArgLysArgProGlyPheGlnHisH 536
Db 5463 GAAATGAGCTGCCTAGACTCATCTATGTTTCTCGTGAAGCGGCTCGAATTTCAACACC 5522
QY 536 isLysLysAlaGlyAlaMetAsnAlaLeuVal----- 546
Db 5523 ACAAAAAGCTGGAGCTATGATGCTATGTTTGTAACTTTCAGAACTTATGTCGCC 5582
QY 546 ----- 546
Db 5583 TCTATTTTATCTCTGTTTCACTGCTTAAGAAACGTTCTCTGTGTAGCCGCTTCCTCA 5642
QY 546 ----- 546
Db 5643 CATTCCTTTTTTCTAGGCTARGTGTCTCTCTAATTTAGTATCTTTTACTTTTGACAG 5702
QY 547 ---ArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHis 565
Db 5703 ATCCGTGTATCTGCTGTTCTTACCATATGAGCATATCTTTTGAACGTGGATGTGATCAT 5762
QY 566 TyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeu 585
Db 5763 TACTTTTAACAGTAAGGCTATTAAAGAAAGCTATGTGTTTCATGATGACCCCGCTATT 5822
QY 586 GlyProGlnValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAsp 605
Db 5823 GGAAGAARGCTGCTATGTCAGTTCCTCAACGTTTTCAGCGTATTGATTTGCACGAT 5882
QY 606 ArgTyrAlaAsnArgAsnThrValPhePheAsp----- 616
Db 5883 CGATATGCCAACAGGAATATAGTCTTTTTCGATGTGAGTATCACTTCCCATTTGTCCTTT 5942
QY 616 ----- 616
Db 5943 GTTTCCTTTTGTTCATATTTTGGTTGATTACTCGTTTCTGCTATGGCCTGACTTGA 6002
QY 617 -----IleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyr 630
Db 6003 TATTTGTTCTCTGGGCAGATTAAACATGAAGGGTTGGATGATATCCAGGGTCCAGTATA 6062
QY 630 rValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleIly 650
Db 6063 TGTGGGTACTGTTGTTGTTTAAATAGCAGGCTCTATATGGGTATGATCCTGTTTGAC 6122
QY 650 sAlaLys-----LysProGlyPheLeuAla---SerLeuCysGlyGlyLysLysAl 667
QY 667

Db 6123 GGAAGAAGATTTTGAACCAAAATATTATTGTCAGAGCTGTTGCGGGTCAAGGAAGAAAGG 6182
QY 667 aSerLysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisValAspSerSe 687
Db 6183 TAAAGTAGTACAGAAAGTAGTATACTAGCAAAAGAGAGAGGCATCAACAGAAAGTGACTCCA 6242
QY 687 rValProValPheAsnLeuGluAspIleGluGlyValGluGly----- 702
Db 6243 TGCTCCACTTTTCAATATATGAGGACATCATGAGGGTTTGAAGGTTTGAATTGAGCTGAT 6302
QY 702 ----- 702
Db 6303 TGTGTAATAACATCATCTTTTATGTAATGATTATGATGATGATGTAATCTTACATCTCT 6362
QY 703 -----AlaGlyPheAspAspGlyLysSerValLeuMetSerGlnMetSerLeuGlyL 720
Db 6363 TGTTTATCAGGTTTATGATGATGAGAGGTCTATTCTTAATGTCCAGAGAGAGTGTAGAA 6422
QY 720 sArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPr 740
Db 6423 GCGTTTGTGTCAGTCGCGGTATTTTATGCGGCAACCTTCATGGGAACAAGCGGCATCC 6482
QY 740 oGlnSerSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGlyTy 760
Db 6483 ACCAAACAACCAATCCCGCTACTCTTCTGAAGGAGGCTATTCTATGTTATAAGCTGTGTTA 6542
QY 760 rLysAspLysSerGluTyrGlyThrGlu----- 769
Db 6543 CGAAGACAAGACTGAATCGGCAAGAGGTCAGTTTTCAAATGCGAGCTACAGAATCTTCT 6602
QY 770 -----IleGlyTyr 772
Db 6603 TATGTTCTTCTTCTACCTGTTTGTGATGATCTTATTTGGCACCTTTTGTAGATTGGTTG 6662
QY 772 pIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyr 792
Db 6663 GATCTATGTTCCGTCGAGCGAAGATATCTTACTGGTTCGAAGTCAAGATGATGCCCGGGTTG 6722
QY 792 pArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLe 812
Db 6723 GATATCGATCTACTGCAATCTCCACGCCCTTCGCTTCAAGGATCTGCACCAATCAATCT 6782
QY 812 uSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSe 832
Db 6783 TTCTGATCGTTTGAACCAAGTCTTTCGATGGCTTTGGATCTATCGAGATCTTCTTAG 6842
QY 832 rArgHisCysProLeuTyrGlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAl 852
Db 6843 CAGACATTTGCTCTCTGTTATGTTTACCATGGAAGGTTGAGACTTTTGGAGAGGATCGC 6902
QY 852 aTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIleIle 872
Db 6903 TTATATCAACACCATCGTCTATCTTATTCATCTCCCTCTTATTGCTGTTATTGTTCT 6962
QY 872 uProAlaIleCysLeuLeuThrGlyLysPheIleMetProGlu----- 886
Db 6963 TCCCGCTTTTGTCTCATCACGAGAGATTCATATACCCGAGGTTTGTAAACACTGACCA 7022
QY 886 ----- 886
Db 7023 CACTGCTATTACTATTTGAATCCCATTTTGTGAATGCAATTTTTTGTATCATCATCTGT 7082
QY 887 -----IleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeuSerIlePheAlat 905
Db 7083 TGCAGATNAGCAACTACCGAGATTTTGGTTCATCTCTACTCTTCACTCAATTTGCTGTGA 7142
QY 905 hrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspLeuTyrTrpArgAsnGluG 925
Db 7143 CTGGAATCCTGGAGCTGAGATGGAGCGGTGTGAGCATTTGAGGATTTGGTGGAGGAACGAGC 7202
QY 925 lnPheTyrValIleGlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuL 945
Db 7203 AGTTCTGGGTCAATGGTGGCACATCCGCCCATCTTTTGTGCTGTCTTCAAGGCTCTACTTA 7262

```

QY 945 ysValLeuAlaGlyIleAspThrAsnPhenThrValThrSerLysAlaAsnAspGluGluG 965
Db 7263 AGGTTCCTGCTGGTATCGACCAACATTCACCGTTACATCTAAAGCCACAGCAAGATG 7322
QY 965 lyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuLeuProProThrThrI 985
Db 7323 GGGATTTCGAGAACTCTACATCTCAATGGACAGCTCTCTCATTTCCACCAACACCG 7382
QY 985 leLeuIlelleAsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyT 1005
Db 7383 TCCTACTGTGTGAACCTCATAGGATTTGGCTGGTGTCTCTTATGCTGTAAACAGTGGCT 7442
QY 1005 yrglnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisL 1025
Db 7443 ACCAGTCGTGGGTCCTTCCTTCGGGAAGCTCTCTTCGCGCTTATGGGTATTGCCCATC 7502
QY 1025 euTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleV 1045
Db 7503 TCTACCTCTTCTGAAAGGCTCTGTGGAGAGACAAAACCGNACCAACCATCGTCAATG 7562
QY 1045 alTrpAlaValLeuAlaSerIlePheSerLeuLeuTrpValArgValAspProPheT 1065
Db 7563 TCTGGTCTGTCTTCGCGCTCCATCTCTCGTTGCTTGGTTCAGGATCAATCCCTTTG 7622
QY 1065 hrThrArgLeuAlaGlyProAsn 1072
Db 7623 TGGAC-----GCCATCCCAAT 7639

RESULT 6
US-08-960-048-1
; Sequence 1, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; PRIORITY FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/029,987
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3328
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-08-960-048-1

Alignment Scores:
Pred. No.: 0 Length: 3328
Score: 3477.00 Matches: 657
Percent Similarity: 73.60% Conservative: 129
Best Local Similarity: 61.52% Mismatches: 176
Query Match: 60.18% Indels: 106
DB: 3 Gaps: 15

US-09-900-237A-30 (1-1080) x US-08-960-048-1 (1-3328)
QY 18 valCysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAla 37
Db 200 GTTTGCCACACTTGTGTGACATGTGTGGTGAATGTTAATGTGAACCTTTTGTGGCT 259
QY 38 CysAspValCysArgPheProValCysArgProCysTrpGluHisGluArgLysGluGly 57
Db 260 TGCATGAATGTAATTTCCCTATTGTGAAGGTTGTTTGGATGATGATCTTAAGGAAGGA 319
QY 58 ThrGlnAlaCysLeuGlnCysLysThrLysTrpLysArgHisArgGlySerProAlaIle 77
Db 320 CGAAAAGCTTGTGGCTGTGT-----GGTAGTCCATAT--- 352

```

```

QY 78 ArgGlyGluGluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAla 97
Db 353 -----GATGAAACCTTGTGGACGAT-----GTCCAG 379
QY 98 SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsn 117
Db 380 AAGCCACCGCGCATCATCGACATCGCTGCA-----CAT 415
QY 118 ThrGlySerGlyAsnValGly---HisProLysTyrAspSerGlyGluIleGlyLeu 136
Db 416 TTGAACAAGTCTCAGGATGTTGAATTCATGCAAGACATATCAGC-----AGTGTG 466
QY 137 SerLysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln 156
Db 467 TCTACATGATGATG-----481
QY 157 MetSerGlyGluIleProGlyAlaSerProAspHisMetMetSerProThrGlyAsn 176
Db 481 -----481
QY 177 IleSerArgArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe 196
Db 482 -----GAAATG 487
QY 197 SerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGln 216
Db 488 GCTGAAGACAAATGGGAATTCGATTTGGAAGAACACAGGTTGGAAGAAAGAAAG 547
QY 217 AspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAla 236
Db 548 AACAGNAGAAGAGAGCTGCAACACT-----AAGGTT 580
QY 237 AlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGlu 256
Db 581 GAAAGAGAGGCTGAAATCCCACTGAGCAACAATAAGAGATAAATGCGCACCGCATGCT 640
QY 257 ThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArg 276
Db 641 TCC---CAGCCCTCTCGACTATAATTCCTCCAAATCCGAAAGACGACTTGACCAATCCGA 697
QY 277 MetValIleValLeuArgLeuValLeuSerIlePheLeuHisTyrArgLeuThrAsn 296
Db 698 ACCGTGATCATATGCGATGATGATCTTGGTCTTTCTTCCATTATCGAGTAACAAAC 757
QY 297 ProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAla 316
Db 758 CCGCTTGACAGTGCTTTTGGACTGTGGCTCACTTCAGTCATATGTGAAATCTGGTTTGA 817
QY 317 LeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeu 336
Db 818 TTTTCTCGGTGTGGATCAGTTCCCTAAGTGTATCTCTGTAAACAGGAAACATAGATT 877
QY 337 AspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAsp 356
Db 878 GACAGACTATCTGCAAGATATGAAGAGAGAGTGAACCTGATGAACCTTGTCAGATTGAC 937
QY 357 IlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThrVal 376
Db 938 TTCTTCGTGAGTACAGTGGATCCATTGAAAGAGCGCTCCATTGATTACTGCCAATACTGTG 997
QY 377 LeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAsp 396
Db 998 CTTTCCATCTCTGCTTGGACTACCGGTGGATAAGGTCTCTTGTATATATCTGATGAT 1057
QY 397 GlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrp 416
Db 1058 GTTGGCGCCATGCTGACATTTGATCTCTAGTAGAAGACACCGCACCTTTCAGAAAGATGG 1117
QY 417 ValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGln 436
Db 1118 GTTCCATTTCTCAAAAAAATTTCCATTGAACCCCGGCGCACCTGAGTTTTACTTCTCAG 1177

```

QY 437 LysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMet 456
Db 1178 AAGATTGATTACTTGAAGATATAAGTGGAGCCCTCTTTTGTAAAGACCTAGACTATG 1237
QY 457 LysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLys 476
Db 1238 AAAAGAGATTATGAAGAGTACAAAATTCGAATCAATGCTTTAGTTGCAAGAGCTCAGAAA 1297
QY 477 ValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnThrArg 496
Db 1298 ACACCTGATGAAGAGTACCAATGCAAGATGAACTCTTGGCCAGGAAATACCCGGCT 1357
QY 497 AspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGly 516
Db 1358 GATCACCTGGCATGATTACAGTTTCTTGGATATAGTGGTCTCGTGACATCGAAGA 1417
QY 517 AsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHis 536
Db 1418 AATGAACCTTCTCGACTCGTTTACGCTCTAGAGAGAAGACCTGGCTACCAACACCAC 1477
QY 537 LysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGln 556
Db 1478 AAAAGGCTGGGCTGAAATGCTTTGGTTAGGGTGCTGCAGTCTTACAAATGCTCCC 1537
QY 557 TyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAla 576
Db 1538 TTCATCCTCAATCTTGATTGTGACCACTATGTTAAACAATAGCAAGCAGTTAGGGAGCA 1597
QY 577 MetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGln 596
Db 1598 ATGTGCTTCTTGATGGACCCCAAGTGTGTCAGATGATGCTATGTGCAGTTTCTCCTCAA 1657
QY 597 ArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAsp 616
Db 1658 AGATTGATGGCATAGATAGGAGTGCATGATGCAATGCAATAGGAACACAGTTTCTTTGAT 1717
QY 617 IleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysVal 636
Db 1718 GTTAACATGAAGGCTTGATGGAATCCAAAGGCCAGTTTATGTGGGAACAGTTGTGT 1777
QY 637 PheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPhe 656
Db 1778 TTCAATAGCAAGACACTTATGGCTATGCTCCACCTTCAATGCCAAGTTTCCCAAGTCA 1837
QY 657 LeuAlaSerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAsp 676
Db 1838 TCCTCTCATCTGTCTGTGTGTCGCCGGCAGAGAACCTTAAGATCCATCAGAG 1897
QY 677 LysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIle 696
Db 1898 CTTTATAGGATGCAAAACGGGAAGAACTTGATGCTGCCACTTTTAACTTAGGGAAT 1957
QY 697 GluGluGlyValGluGlyAlaGlyPheAspAsp---GluLysSerValLeuMetSerGln 715
Db 1958 GAC-----AATTATGATGATGATGAAGATCAATGTTGATCTCTCAA 1999
QY 716 MetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGlu 735
Db 2000 ACAAGCTTGAGAAACTTTTGGCTATCTTCAGCTCTTCATGGAATCTACACTAATAGAG 2059
QY 736 TyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisVal 755
Db 2060 AATGAGGAGTGGCTGAATCTGCCAACCCCTCCACACTAATCAAGGAAGCAATTCATGTC 2119
QY 756 IleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGly 775
Db 2120 ATCACTGTGGCTATGAAGAGAAGACTGCATGGGGAAGAGATTGGATGGATATATGT 2179
QY 776 SerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerVal 795
Db 2180 TCAGTCACTGAGGATATCTTAACCGCTTCAAAATGCATCGCCGAGGATGGATCGATT 2239
QY 796 TyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArg 815

Db 2240 TACTGCATGCCCTTAAGGCCAGCAATTCAGAGGATCTGCACCCATCAATCTGTCTCATCGG 2299
QY 816 LeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCys 835
Db 2300 TTGCACAGGTTCTTCGATGGGCTCTTGAAATCTGTGTAAATTTTCTTAAGCAGGATGTC 2359
QY 836 ProLeuTyrTyrGlyTyr---GlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIle 854
Db 2360 CCTCATGGTATGGCTTTGGAGGTGCTGTCTTAATGGCTTCAAGAGACTAGCATATATA 2419
QY 855 AsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuProAla 874
Db 2420 AACACCATTTGCTATCCTTTCCATCCTCTCCACTCATTTGCTATTTGTTCACTACAGCA 2479
QY 875 IleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyr 894
Db 2480 ATCTGCTTCTCACAGGAAATTTATCATCAACAGCTCTCAACCTCGCAAGTGTCTTC 2539
QY 895 PheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGly 914
Db 2540 TTTCTTGCCCTTTTCTTTCCATTATCGTACTGCTGTCTCGAGCTCGATGGAGTGT 2599
QY 915 ValGlyIleAspGluTyrTyrArgAsnGluGlnPheThrValIleGlyGlyIleSerAla 934
Db 2600 GTCAGCATTTAGGACTTATGGCGTAAACGAGCAGTTTGGGTCTATCGGTGGCGTTTCAGCC 2659
QY 935 HisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPhe 954
Db 2660 CATCTCTTTGGCGTCTTCCAAAGTTTCTTAAGATGCTTGGGGCATTTGACACCAACTTT 2719
QY 955 ThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLys 974
Db 2720 ACTGTCACCTGCCAAGCAGCTGAT---GATGCAGATTTTGGTGACCTCTACATTGTGAAA 2776
QY 975 TrpThrThrLeuLeuLeuProProThrThrIleLeuIleIleAsnMetValGlyValVal 994
Db 2777 TGGACTACATCTTAATCCCTCCAAACACACTCTCATCGTCAACATGGTTGGTCTGCTT 2836
QY 995 AlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLys 1014
Db 2837 GCGCGATTCTCCGATGCCCTCAACAAAGGTACGAAGCTTGGGGACCACTCTTTGGCAAA 2896
QY 1015 LeuPhePheAlaPheTyrValIleValHisLeuTyrProPheLeuLysGlyLeuMetGly 1034
Db 2897 GTGTCTTTCTTCTTGGGTCTATCCTCATCTTATCCATTCCTCAAGGCTCTTATGGGA 2956
QY 1035 ArgGlnAsnArgThrProThrIleValIleValTyrAlaValLeuLeuAlaSerIlePhe 1054
Db 2957 CGCCAAAACAGGACACCAACCACTTGTCTCTTGGTCAGTGTGTGGTCTTCTGCTCTTC 3016
QY 1055 SerLeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle--- 1073
Db 3017 TCTCTTGTGGTTCGGATCAACCCGTTTGTACACCGCGGATAGCACCACCGGTGCA 3076
QY 1074 GlnThrCys---GlyIleAsnCys 1080
Db 3077 CAGAGCTGCATTTCCATTGATTGT 3100

RESULT 7

US-09-838-586-1
; Sequence 1, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Straker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048

; PRIOR FILING DATE: 1997-10-29
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3328
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic oligonucleotide
 US-09-838-586-1

Alignment Scores:

Pred. No.: 0 Length: 3328
 Score: 3477.00 Matches: 657
 Percent Similarity: 73.60% Conservative: 129
 Best Local Similarity: 61.52% Mismatches: 176
 Query Match: 60.18% Indels: 106
 DB: 4 Gaps: 15

US-09-900-237a-30 (1-1080) x US-09-838-586-1 (1-3328)

Qy 18 ValCysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAla 37
 Db 200 GTTTGCCACACTTGTGGTGAACATGTTGGGTGAATGTTAATGTTGAACCTTTTGTGGCT 259
 Qy 38 CysAspValCysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGly 57
 Db 260 TGGCATGAATGTAATTTCCCTATTGTAAGAGTTGTTTGGATGATGATCTTAAGGAAGGA 319
 Qy 58 ThrGlnAlaCysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIle 77
 Db 320 CGAAAGCTTGTTCGCTTGT-----GGTAGTCCATAT--- 352
 Qy 78 ArgGlyGluGluGlyAspThrAspAlaAspGlySerAspPheAsnTyrProAla 97
 Db 353 -----GATGAAACCTGTTGGACGAT-----GTCGAG 379
 Qy 98 SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsn 117
 Db 380 AAGGCCACCGCGCATCAATCGACNAATGGCTGCA-----CAT 415
 Qy 118 ThrGlyGlySerGlyAsnValGly---HisProLysTyrAspSerGlyGluIleGlyLeu 136
 Db 416 TTGACACAGTCTCAGGATGTTGGAAATTCAGCAGACATATCAGC-----AGTGTG 466
 Qy 137 SerLysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln 156
 Db 467 TCTACATTGGATAGT----- 481
 Qy 157 MetSerGlyGluIleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsn 176
 Db 481 ----- 481
 Qy 177 IleSerArgArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe 196
 Db 482 -----GAAATG 487
 Qy 197 SerGlySerIleGlyAsnValAlaTyrPlysGluArgValAspGlyTyrPlysMetLysGln 216
 Db 488 GCTGAACACCAATGGGAATTCGATTGGAAGAACACAGGCTGGAAGTTGGAAGAAAGAAAG 547
 Qy 217 AspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAla 236
 Db 548 AACAAAGAAAGAGAGCTGCAACAACT-----ARGGTT 580
 Qy 237 AlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGlu 256
 Db 581 GAAGAGAGGCTGAATCCCACTGAGCAACAATGGAAGATAAACCGGCACCGGATGCT 640
 Qy 257 ThrArgGlnProLeuSerArgLysValProIleAlaSerSerIleAsnProTyrArg 276
 Db 641 TCC---CAGGCCCTCTCGATATATTCCATCCCGAAAGACAGACTTGCACCATACCGA 697

Qy 277 MetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsn 296
 Db 698 ACCGTGATCAATATGCGATGATCAATCTTGGTCTTTTCTTCCATTAATCAGATTAACAACAA 757
 Qy 297 ProValArgAsnAlaTyrProLeuTyrLeuSerValIleCysGluIleTyrPheAla 316
 Db 758 CCCGTTGACAGTGTCTTTGGACTGTGGCTCAGTCATATGTGAATCTGGTTGCA 817
 Qy 317 LeuSerTyrIleLeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLeu 336
 Db 818 TTTTCTCGGTGTGGATCAGTTCCCTAAGTGTATCCTGTAAACAGGAAACATACATAT 877
 Qy 337 AspArgLeuAlaLeuArgTyrAspArgGluGluProSerGlnLeuAlaAlaValAsp 356
 Db 878 GACAGACTCTCGAAGATATGAAGAAGAGGTGAACCTGATGAACCTGTGAGTTGAC 937
 Qy 357 IlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThrVal 376
 Db 938 TTCTTCGTGAGTACAGTGGATCCATTGAAGAGCCTCCATTGATTACTGCCAATACTGTG 997
 Qy 377 LeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAsp 396
 Db 998 CTTTCCATCTCTGCTTGGACTACCCGCTGGATAAGGTCTCTTGTATATATCTGATGAT 1057
 Qy 397 GlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrp 416
 Db 1058 GGTGCGGCCCATGCTGATTTGAATCTCTAGTAGAAACACGCCACTTGTCAAGAAGTGG 1117
 Qy 417 ValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGln 436
 Db 1118 GTTCCATCTCTGCAAAAATTTTCCATTGAACCCCGGCACCTGAGTTTACTTCTCAG 1177
 Qy 437 LysIleAspTyrIleLysAspLysValGlnProSerPheValLysAspArgAlaMet 456
 Db 1178 AAGATTGATTACTTGAAGATATAAGTCAGCCCTCTTTGTAAAGAACCTGAGACTATG 1237
 Qy 457 LysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLys 476
 Db 1238 AAAAGAGATTATGAAGAGATCAAAATTCGAATCAATCTTTAGTTGCAAGGCTCAGAAA 1297
 Qy 477 ValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArg 496
 Db 1298 ACACCTGATGAAGATGACATGCAAGATGGAATCTTTGGCCAGGAATAACCCCGT 1357
 Qy 497 AspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGly 516
 Db 1358 GATCACCTCGCATGATTCAAGTTTCTCTGGATATAGTGTGCTCGTCGATCGAAGGA 1417
 Qy 517 AsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHis 536
 Db 1418 AATGAATCTCTCGACTGGTTTACGCTCTAGAGAGAGAGACCTGCTACCAACACCAAC 1477
 Qy 537 LysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGln 556
 Db 1478 AAAAAGCTCGTGTGTAATAATGCTTTGTTAGGTGTCTGAGTTCTTACAAATGCTCCC 1537
 Qy 557 TyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAla 576
 Db 1538 TTTCACTCTCAATCTTGATTGTGACCATATGTTAAACAATAGCAAGGAGGTTAGGAGGCA 1597
 Qy 577 MetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGln 596
 Db 1598 ATGTGCTTCTTGATGGACCCCAAGTTGGTCGAGATGTATGCTATGTCAGTTTCTCTCAA 1657
 Qy 597 ArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAsp 616
 Db 1658 AGATTTGATGGCATAGTAGAGTATGATATGCCAATAGGAACACAGTTTCTTTGAT 1717
 Qy 617 IleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysVal 636
 Db 1718 GTTAAACATGAAGGTCTTGATGGAATCCAAAGGCCAGTTTATGTGGAAACAGGTTGTGT 1777
 Qy 637 PheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPhe 656

Db 1778 TTCAATAGCGACGACCTTTATGGCTATGGTCCACCTTCAATGCCAAGTTTCCCAAGTCA 1837
Qy 657 LeuAlaSerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAsp 676
Db 1838 TCCTCTCATCTTGTCTGTCTGTCTGCGCGGCAAGAAGAACCTTAAAGATCCATCAGAG 1897
Qy 677 LysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIle 696
Db 1898 CTTTATAGGATGCAAAACGGAAGAACTTGATGCTGCCATCTTTAACTTAGGGAAT 1957
Qy 697 GluGluGlyValGluGlyAlaGlyPheAspAsp---GluLysSerValLeuMetSerGln 715
Db 1958 GAC-----AATTATGATGATGATGAAGATCAATGTTGATCTCTCAA 1999
Qy 716 MetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGlu 735
Db 2000 ACAAGCTTTGAGAAACTTTTGGCTTATCTTCAGTCTTCATTGAATCTACACTAATGGAG 2059
Qy 736 TyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisVal 755
Db 2060 AATGAGAGTGGCTGAATCTGCCAACCTTCCACACTATCAAGAGCAATTCATGTC 2119
Qy 756 IleSerCysGlyTyrCysGlyLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGly 775
Db 2120 ATCAGCTGTGGCTATGAAGAGAGACTGCATGGGGGAAAGAGATTGGATGATATATGCT 2179
Qy 776 SerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerVal 795
Db 2180 TCAGTCATCTGAGGATATCTTAACCGGCTTCAAAATGCACTGCCGAGGATGGAGATCGATT 2239
Qy 796 TyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArg 815
Db 2240 TACTGCATGCCCTTAAGCCACATTCAAAGATCTGCACCCATCAATCTGTCATCGG 2299
Qy 816 LeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCys 835
Db 2300 TTGCACCAAGTTCTTCGATGGGCTCTTGGATCTGTGTGAATTTCTTAAGCAGGCATTGC 2359
Qy 836 ProLeuTrpTyrGlyTyr---GlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIle 854
Db 2360 CCTCATGTATGGCTTATGGAGGTGGTCTGTCTTAATGGCTTCAAGAGCTAGCATATATA 2419
Qy 855 AsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuProAla 874
Db 2420 AACACCATGTCTATCTTTCATCTCCCTTCCACTCATTCGCTATGTTCACTACCAGCA 2479
Qy 875 IleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTrp 894
Db 2480 ATCTGTCTTCTCACAGGAAATTTATCATACCAGCTCTCAAACTGGCAAGTGTCTTC 2539
Qy 895 PheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGly 914
Db 2540 TTTCTTGGCCCTTTCTTCTTCCATATATCGTACTGTCTGCTCTCGAGCTCCGATGGAGTGT 2599
Qy 915 ValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyIleSerAla 934
Db 2600 GTACGATTTAGGACTTATGGCGTAACGAGCAGTTTGGGTCTATCGGTGGCGTTCAGCC 2659
Qy 935 HisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPhe 954
Db 2660 CATCTCTTGGCGTCTTCCAGGTTCCTTAAGATGCTTGGCGGCAATTGACACCACTTT 2719
Qy 955 ThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLys 974
Db 2720 ACTGTCACGCAAGACGCTGAT---GATGAGATTTTGGGAGCTCTACATTTGGAAA 2776
Qy 975 TrpThrThrLeuLeuIleProThrThrIleLeuIleIleAsnMetValGlyValVal 994
Db 2777 TGGACTACACTTCTAATCCCTCCAAACACACTCTCATCTGTCACATGTTGGTGTGCT 2836
Qy 995 AlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLys 1014

Db 2837 GCCGATTTCTCCGATSCCTCAACAAGGGTACGAAGCTTGGGACCACCTCTTTGGCAAA 2896
Qy 1015 LeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGly 1034
Db 2897 GTGTTCTTTTCTCTCTGGGTCACTCCATCTTTATCCATCTCTCAAGGCTTTATGGGA 2956
Qy 1035 ArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePhe 1054
Db 2957 CGCCAAACAGACACCAACCATTTGTTGCTTCTGTCAGTGTGTTGGCTTCTGTCITC 3016
Qy 1055 SerLeuLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle--- 1073
Db 3017 TCTCTTGTGTTGGTTCGGATCAACCCGTTTGTTCAGCACCGCGCATAGCACCCGCTGTCA 3076
Qy 1074 GlnThrCys---GlyIleAsnCys 1080
Db 3077 CAGAGCTGCATTTCCATTGATGT 3100
RESULT 8
US-09-221-013A-1
; Sequence 1, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2248
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1887)
US-09-221-013A-1
Alignment Scores:
Pred. No.: 2,36e-273 Length: 2248
Score: 2629.50 Matches: 468
Percent Similarity: 87.46% Conservative: 76
Best Local Similarity: 75.24% Mismatches: 71
Query Match: 45.51% Indels: 7
DB: 4 Gaps: 4
US-09-900-237A-30 (1-1080) x US-09-221-013A-1 (1-2248)
Qy 454 ArgAlaMetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLys 473
Db 1 CGAGCTATGAGAGAGAGATGATGAAGAGTTTAAAGTGAGATAAATGCTCTTGTGGCAAA 60
Qy 474 AlaLeuLysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsn 493
Db 61 GCACAGAAAATCCCTGGAGAGGCTGGACAATGAGGATGGTACTCCCTCGGCTGGTAAC 120
Qy 494 AsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyIleAsp 513
Db 121 AACACTAGAGATCATCTCGGAATGATACAGGTGTTCTTAGGCCATAGTGGGGTCTGGAT 180
Qy 514 ThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPhe 533
Db 181 ACCGATGGAATGAGCTGCCTAGACTCATCTATGTTCTCGTGAAGAGCGGCTGGATTT 240
Qy 534 GlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThr 553

Db 241 CAACACACAAAAGGCTGGAGCTATGAATGCATCGATCGGTATCTGCTGTTCTTACC 300
QY 554 AsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaVal 573
Db 301 AATGGAGCATATCTTTTGAACGTTGGATGTCATTACTTAAATACAGTAAGGCTATT 360
QY 574 ArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGln 593
Db 361 AAGAAGCTATGTTTTCATGATGGACCCCGCTATTGGAAGAAGTCTGCTATGTCAG 420
QY 594 PheProGlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrVal 613
Db 421 TTCCTCAACGTTTGAAGGTTATGATTTGCACGATCGATATGCACAGGAATATAGTC 480
QY 614 PhePheAspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThr 633
Db 481 TTTTTCGATATTAAATCAAGAGGGTGGATGGTATCCACGGTCCAGTATATGGGTACT 540
QY 634 GlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLys--- 652
Db 541 GGTGTGTTGTTTAAATAGCGAGCTCTATATGGGTATGATCCCTGTTTTCGCGAAGAAGAT 600
QY 653 ---LysProGlyPheLeuAla---SerLeuCysGlyGlyLysLysLysAlaSerLysSer 670
Db 601 TTAGAACCAATATATTGTCAAGAGCTGTTCGGGTCAAGGAAGAAGGTTAAAGTAGC 660
QY 671 LysLysArgSerSerAspLysLysLysSerAsnLysHisValAspSerSerValProVal 690
Db 661 AAGAAGTATATACTACGAAAAGAGAGGAGCATCAACAGAGTCACTCCAACTGCTCCACTT 720
QY 691 PheAsnLeuGluAspIleGluGluGlyValGluGlyAlaGlyPheAspAspGlyLysSer 710
Db 721 TTCATATGAGGACATCGATGAGGGTTTGA-----GGTTATGATGATGAGAGGTCT 774
QY 711 ValLeuMetSerGlnMetSerLeuGlnLysArgPheGlyGlnSerAlaAlaPheValAla 730
Db 775 ATTCATATGCCAGAGAGGTAGAGAGGTTTGGTCAGTCGCCGGTATTATTGGC 834
QY 731 SerThrLeuMetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLys 750
Db 835 GCACCTTCATGGAACAGGCGGCATTCACCAACCAACCACTCCGCTACTCTTCTGAAG 894
QY 751 GluAlaIleHisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIle 770
Db 895 GAGGCTATTATGTTATAGCTGGTTACAGACAGCAAGTGAATGGGCAAAAGAGATT 954
QY 771 GlyTyrIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArg 790
Db 955 GGTGGATCTATGTTCCGTGACGAAGATATTCTTACTGGTTCAAGATCATGCCCGG 1014
QY 791 GlyTyrArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIle 810
Db 1015 GGTGGATATCGATCTACTGCAATCCTCCACCGCTCGTTTCAAGGGATCTGCACCAATC 1074
QY 811 AsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeu 830
Db 1075 AATCTTCTGATCGTTGAAACCAAGTCTTCGATGGGCTTTGGGATCTATCGAGATTCTT 1134
QY 831 PheSerArgHisCysProLeuTyrPheGlyTyrGlyArgLeuLysPheLeuGluArg 850
Db 1135 CTTAGCAGACATTGCTCTATCTGGTATGTTACCATGGAAGTTGAGACTTTTGAGAGG 1194
QY 851 PheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCys 870
Db 1195 ATCGCTTATATCAACACCATCGTCTATCTCTATTATCATCCATCCCTCTATTGCGTATTGT 1254
QY 871 IleLeuProAlaIleCysLeuThrGlyLysPheIleMetProGluIleSerAsnLeu 890
Db 1255 ATTCCTCCCGCTTTTGTCTCATCCACGACAGATTATCATACCGGAGATGAAGCACTAC 1314
QY 891 AlaSerIleTyrPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMet 910

Db 1315 GCGAGTATTGGTTTCATTCTACTCTTTCATCTCAATTGCTGTGACTGGAATCCTGAAACTG 1374
QY 911 ArgTyrSerGlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIleGly 930
Db 1375 AAATGGAAACGCTGTGAGCATTTGAGATTGGTGGAGGAACAACCAAGTTCTGGGTCAATTGGT 1434
QY 931 GlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLysValLeuAlaGlyIle 950
Db 1435 GGCATATCCACCATCTTTTGTGCTCTTCCAGGTCTTCTTAAGGTCTTGTCTGGTATC 1494
QY 951 AspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeu 970
Db 1495 AACACCAACTTCACCGTTTACATCTTAAAGCCACAAACAAATGGGATTTTGCAAACTC 1554
QY 971 TyrMetPheLysTyrThrThrLeuLeuLeuProThrThrIleLeuIleAsnMet 990
Db 1555 TACATCTCAATGGACAGCTCTTCTCATTCACCAACCAACCGCTCTACTTGTGAACCTC 1614
QY 991 ValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyPro 1010
Db 1615 ATAGGCAATTTGGCTGGTGTCTCTTATGCTGTAAACAGTGGCTACCACTGCTGGGGTCCG 1674
QY 1011 LeuPheGlyLysLeuPheAlaPheTrpValIleValHisLeuTyrProPheLeuLys 1030
Db 1675 CTTTTCGGGAAGCTCTTCTTCGCTTATGGGTTATTGCCATCTCTACCCCTTCTTGAAA 1734
QY 1031 GlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeu 1050
Db 1735 GGTCTGTGGAGAGACAAACCGAACCAACCACTGCTCATTTGTCTGTTCTTCTCTC 1794
QY 1051 AlaSerIlePheSerLeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGly 1070
Db 1795 GCCTCCATCTTCGTGCTTTGGGTCAGATCAATCCCTTTGTGGAC-----GCCAAT 1848
QY 1071 ProAsn 1072
Db 1849 CCAAT 1854
RESULT 9
US-09-221-013A-4
; Sequence 4, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 5009
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(5009)
; OTHER INFORMATION: N is A, T, G or C.
US-09-221-013A-4
Alignment Scores:
Pred. No.: 5.25e-263 Length: 5009
Score: 2539.50 Matches: 559
Percent Similarity: 56.71% Conservative: 130
Best Local Similarity: 46.01% Mismatches: 166
Query Match: 43.95% Indels: 362


```
Db 3491 ATTGTTGTAATACACAGGTGTTCTTTGGGTCAATAGTGGAGTTTCGTGATACGGAGTGTAAATGA 3550
Qy 518 uLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLy 538
Db 3551 GTTACCACGCTCTAGTGTATGTTCTCGTGAGAGCGGCTGGATTGATCACCAAGAA 3610
Qy 538 sAlaGlyAlaMetAsnAlaLeuVal----- 546
Db 3611 AGCTGGAGCTATGAATTCCTTGTTGTAAGTATAATGTGTTCTTTATTTATGATCTCTCTT 3670
Qy 547 -----ArgVa 548
Db 3671 TTCGGAGCCCTGACTTCTCATATAAACTAAACTCATCTTACTTCTTCTTGAAGATCCGAGT 3730
Qy 548 lSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAs 568
Db 3731 CTCTGCTGTCTTATCAAAACGCTCTTACCTTCTTAATGTGCAATGTGATCACTACATCAA 3790
Qy 568 nAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyProG 588
Db 3791 CAACAGCAACCAATTAGAGNATCTATGTGTTTCATGATGACCCGCAATCGGGAAGAA 3850
Qy 588 nValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArgTyrAl 608
Db 3851 AGTTTGTATGTTTCAGTTTCGCGACAGATTTGATGGGATTGATAGACATGATAGATACTC 3910
Qy 608 aAsnArgAsnThrValPhePheAsp----- 616
Db 3911 AAACCGTAACGTTGTGTTCTTTGATGATGTCCTTATCTCTTTGCTTTGTTTCGTGTT 3970
Qy 617 -----lleA 618
Db 3971 TATGTTTATAGCTTTTCCCTCTTTCTCAITTCATGATATGTTTGGTGTGAAACAGATTA 4030
Qy 618 snLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheA 638
Db 4031 ACATGAAGGCTCTGATGGGATACAGGACCGATATATGTCGGACAGGTGTGTGTGTTA 4090
Qy 638 snArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuA 658
Db 4091 GAAACAGGCTCTTATGGTTTGTATGCACCAAGAAAGAAACCAACCCAGGCAAAACCT 4150
Qy 658 laSer-----LeuCysGlyGlyLysLysLysAlaSerLysSerL 671
Db 4151 GTAACCTTGGCCCTAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4204
Qy 671 yLysArgSerSerAspLysLysSerAsnLysHisValAspSerSerValProValP 691
Db 4205 AAACGAAAGCCAAAGATAAGAAACTAACACTMAA-----GAGACTTCAAAGCAGATTC 4258
Qy 691 heAsnLeuGluAspIleGluGlyVal-----GluGlyAla--- 703
Db 4259 ATGCGCTAGAGAAATGTCGAGGAAGTGTATGCTCCAGGTATATGCTCCAGGTATAAAAAGGAAAAA 4318
Qy 704 -----GlyPhe----- 705
Db 4319 AAACATTTCTATTGTTGTTCTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTA 4378
Qy 706 -----AspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGly 723
Db 4379 GTGTCAAAATGTTGAGAAGAGATCTGAAGCAACACAAATGAAATGGAGAGAAGTTTGGGA 4438
Qy 724 GlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValProGlnSerSer 743
Db 4439 CAATCTCCGTTTTCGTTGCTCTGCTGTTCTTACAGAACCGTGGAGTTCCTCCGTAACGCA 4498
Qy 744 ThrProGluSerLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLys 763
Db 4499 AGCCCCGATGTTTGTTAAGAGAGACCAATTCAGTTATTAGCTCGGGGTACGAAGATAAA 4558
Qy 764 SerGluTrpGly----- 767
Db 4559 ACCGAATGGGAAAAGAGGTAGAAAAACATTACAAAGTTTTTCAACTTCTGAAACTAGAA 4618
```

```
Qy 768 -----ThrGluIleGlyTrpIleTyrGlySe 776
Db 4619 AAGTTCTTGTGATCTCAATTCCTGCTGATAATCACACGACGATCGGTGGATTTATGGATC 4678
Qy 776 rValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTy 796
Db 4679 GGTGACTGAAGATATCTGACGGGTTTCAAGATGCATTCATGATGGAGATCTGTGTA 4738
Qy 796 rCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLe 816
Db 4739 CTGTATGCCTAAGCGTCAGCTTTTAAAGGATCTGCTCTTAACTATTGTTCAGATCGTCT 4798
Qy 816 uAsnGlnValLeuArgTrpAlaLeuGlySerValGluLeuPheSerArgHisCysPr 836
Db 4799 TCATCAAGTTCTACGTTGGGCTCTTGCTCTGTAGAGATTTTCTTGGACAGACATTTGTCC 4858
Qy 836 oLeuTyrTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnTh 856
Db 4859 GATATGTTATGTTATGTTGTTTAAATGGTTGGAGAGATCTCTTACATCAATC 4918
Qy 856 rThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCy 876
Db 4919 TGTGCTCTATCTTGGACTTCACCTTCATTCATGATGCTCTATTGTTCTCTCCCGGGTTTG 4978
Qy 876 sLeuLeuThrGlyLysPheIleMetProGlu 886
Db 4979 TTTACTCACAGGAAATTCATCGTCCCTGAG 5009
RESULT 10
US-08-960-048-2
; Sequence 2, Application US/089600048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Scalkar, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; TITLE OF INVENTION: Sequences
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; CURRENT FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4612
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-08-960-048-2
Alignment Scores:
Pred. No.: 9,71e-187 Length: 4612
Score: 1833.00 Matches: 510
Percent Similarity: 42.72% Conservative: 80
Best Local Similarity: 36.93% Mismatches: 78
Query Match: 31.72% Indels: 714
DB: 3 Gaps: 29
US-09-900-237A-30 (1-1080) x US-08-960-048-2 (1-4612)
```

```
Qy 413 AlaArgLysTrpValProPheValLysTyr-Asp----- 424
Db 39 GCGAGGAGATGGGTTCCGTTTGTAAAGACATTCACCTAGGGGGCCGACGTCTCTT 98
Qy 425 -----IleGluProArgAl 429
Db 99 AAGCGGTGCTGCTCTCTCTACCAAGCAAAACATTTCTTGTAAATGTTGAGCCCGAGGC 158
Qy 429 aProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGln----- 446
Db 159 GCCGGAGTTTATTTCAATGAGAGAAATTGATTATTATTGAAGGACAAAGTCCATTATCAAC 218
```

446 ----- 446
Db TCGGGTCCGGCGCTCAAAATAAAGTTACTCTCTTAATAATAAACTTCCTGTCACAG 278
447 --ProSerPheValLysAspArgAlaMetLysArgGluTyrGluGluPheLysIleA 466
Db TACCTAGCTTTGTTAAAGAACCGAGAGCCATGAAAGGGAATATGAAGATTTAAAGTAA 338
446 rGileAsnAla----- 469
Db GGATCAATGCATGATCGAACAATTTCTTGCTCTCGGTACTTTCCCTTATACCTTCT 398
470 -----LeuValSerLysAlaLeuLysValProGluGluGlyTyr 482
Db AAAATTCATTCCTAGTTAGTATAGTACCAAAAGCTCAGAGAAACCAAGAGATGG 458
483 IleMetGlnAspGlyThrProTyrProGly-Asn----- 493
Db GTGATGCAAGATGGCACCCCATGGCCCGGAAATCATCGTTTTCGAGTCTCTTTGGTCT 518
494 -----AsnThrArgAspHisPr 499
Db TCTTCTACCCACTAGTTCTACCGTGGGGTACCGGGCTTTTAAACACTCGTATCATCC 578
499 ocLysMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGly- 516
Db TGGATGATTCAGGTCTATCTAGGAAGTGCCTGGTGCATCGATGTGGATGGCAATTGTGAG 638
516 ----- 516
639 CACTAGTAGGACCTTACTAAGTCCAGATAGATCTTACCGGCCACGTGAGCTACACCTAC 698
517 --AsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisH 536
Db CGAAAGAGCTGCCTCGCTGTCTATGTTCTCGTGAGAAACGACCTGGTTATCAGCACC 758
536 isLysLysAla----- 539
Db ATAGAAAGCCGTTTCTCGACGGAGCTGNAACAGATACAAAGAGCACTCTTCTGCTGGACCA 818
540 -----Gly-AlaMetAsnAlaLeuValArgValSerAlaValLeu 552
Db ATAGTCGTGGTATCTTTCGGCGTCTGAGAAATGCTCTGGTTCGAGTTCTCGAGTGTCT 878
553 ThrAsnGlyGlnTyrMetLeuAsnLeuAspCys----- 563
Db ACTAATGCACCTTCATATGTAATCTGGATTGACGACTCTTTACGAGACCAAGCTCAAAG 938
564 -----AspHisTyrIleAsnAs 569
Db AGTCACGNATGATACGTGGGAAGTATATCTTAGACCTAATCTGATCATATCAACAA 998
569 nSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyPro----- 587
Db TAGCAAGGCCATGAGGAAGCATGTGCTTTTAAATGGATCCTCAGTTTGGAACTAGTAA 1058
587 ----- 587
1059 TGATGTGTTATCGTTCCGGTACTCCCTTCGCTACACGAAAAATTTACCTAGGAGTCAAAAC 1118
588 -----Gln-ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspA 606
Db CTAAGAGCTTTGTTATGTTCAATTTCCACAGAGATTTGATGGTATTCATGATC 1178
606 rGlyTyrAlaAsn----- 609
Db GATATGCTAATCTTCTTCAAAACAATAACAAGTTAAAGGTGTCTCTAAACATACCATAACTA 1238
610 -----Arg-AsnThrValPhePheAspIleAsnLeuArgGlyLeu 622
Db GCAGTACTAGCTATACGATTAGGAATGTTGTTCTTTTATGATATCAACATGTTGGGATTA 1298

623 AspGlyIleGlnGlyProValTyrValGlyThr----- 633
Db GATGGACTTCAAGGCCCTGTATATGTAGGCACCTTTACAACAGAAAAAATACTAGTTGA 1358
634 -----GlyCysValPheAsnAr 639
1359 CAACCTAATCTACCTGAAGTTCGGGACATATACATCCGTGAGGCTGTGTTTCAACAG 1418
639 gThrAlaIleTyrGlyTyrGluProPheIleLysAlaLysLysPro----- 654
Db CGAGCATTTGATGGCTACGATCCACAGTCTCTGAGAAACGACCAAGAGATGCCACAC 1478
654 ----- 654
1479 AAAAGTTCTCGCTCCGTAAACATACCGATGTAGTGGTTCAGAGACTCTTTGCTGTTTCT 1538
655 -----GlyPheLeuAla---SerLeu-CysGlyGlyLysLysLysAlaAs 668
Db ACACATGTGATGTGCGCTTCTTGGTGTCTGTTGTGGAGGTTCTAGGAAGAAAT 1598
668 erLysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisVal----- 684
Db CAAGAAGAAAGTGTACACTAACGACCGAGAACCAACACACACAA-CGCTTCCAAAG 1657
684 ----- 684
1658 ATCCTTCTTTAGTTCTTCTTTCGTGAAAAAGAGGCTTACTCGGAGGTCTTTTATACGG 1717
684 ----- 684
1718 AAAAAAGAAAGATGATGGGCAAAACATATGTCATCTTTTCTCCGAATGAGCCTCCAG 1777
685 -----AspSerSerValP 689
1778 AAAATATGCTTTTCTTCTTCTACTACCGCTTTTGTATACAGAAAAAGGCTCTGCAC 1837
689 toValPheAsnLeuGluAspIleGluGluGlyValGluGly----- 702
Db CAGTCTTTGATCTCGAAGAAATCGAAGAGGCTTGAAGGATACGAAGAATTTGCTTTT 1897
703 -----AlaGlyPheAsp- 706
1898 CCCAGACGTGGTFCAGAACTAGAGTCTTTTAGCTTCTTCCGAACTCTCTATGCTTCT 1957
707 AspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAla 726
1958 AACGAGAAATCCACATTATGTCGCAAGAAATTCGAGAAACGATTCGACAAATCACCG 2017
727 AlapheValAlaSer----- 731
2018 GTTTTCAVTCCTCTCTTTTAGCTGTAATTTACAGCGTCTTCTTAAAGCTCTTTGCTAAGCC 2077
732 -----ThrLeuMetGluTyrGlyGlyValProGlnSerSe 743
2078 TCTTAGTGCCAAAAAGTAAACGACAACTTTGATGAAATGTTGGCTTCTCTGAAGAAAC 2137
743 rThrProGluSerLeuLeuLysGluAlaIleHis----- 754
2138 TAAATCCACATCACTGATTTAAAGAGGCCATTCAGTTGAAACTACCTTTTACCACCGAAG 2197
755 -----ValIleSerCysGlyT 760
2198 GACTTCTTGTAAAGGTGTAGTACTAATTTCTCCGTAAGTCGTAATTTAGCTGTGGT 2257
760 rGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerVal----- 777
2258 ATGAGAAAAAACTGAGTGGGCAAAAGAGATCGATGATTTATGCTCGTGGCAATTAA 2317
777 ----- 777
2318 TCGACACCAATATCTTTTGTGACTCACCCGTTTCTCTAGCTACTTAATATACCCAGC 2377
778 ---ThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrTrpArgSerValTyr 796

```
Db 2378 CACACGGAAGATATATTAAAGGTTTCAAGATGCAATGTAGAGGGTGAATCGTTTAT 2437
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 797 CysMetProLysArg----- 801
Db 2438 TGTGTACCGAAATGCTCTATATAATTGTCAAAGTTCTACGTAACATCTCCACCTT 2497
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 802 -----ProAlaPheLysGlySerAlaProIleAsnLeuSe 813
Db 2498 TAGCCAAATTAACATCGCTTTTACCGGCANTCAAAGGGTCCGCTCCAATCAATCTCTC 2557
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 813 rAspArgLeuAsnGlnValLeuArgTrpAlaLeu----- 824
Db 2558 GGATCGGTTCACCAAGTTTIGAGATGGCACCTCTGGCCGTAAGTTTCCAGCGGAGGTT 2617
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 825 -----GlySerValGluLeuL 830
Db 2618 AGTTAGAGAGCTAGCAACAGTGTTCAAACTCTACCCGCTGATGGTTCTGTAGAAATTT 2677
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 830 euPheSerArgHisCysProLeuTrpTyrglyTyrglyGlyArgLeuLys----- 846
Db 2678 TCCTTAGTCGCTACTGTCCACTTTGGTATGTTAAGTGGAAACTGAAATGGACCAAGA 2737
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 847 -----P 847
Db 2738 CATCTTTAAAGGAATCAGCAGTGACAGGTGAACACCATACCAATACCACCTTTTGACTTT 2797
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 847 he-LeuGluArgPheAlaTyrlleAsnThrTrpIleTyrrProLeuThrSerLeuProLeu 866
Db 2798 ACCCTCGAGAGGCTGTGTTATATCAACACCATTTGTTACCCCTTTTCACTCGATCCCTTTA 2857
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 867 LeuValTyrrCys----- 870
Db 2858 CTCGCTATTGTAGAGCTCTCCGAAACGAATAGTTGTGTGAACAAATGGGAAGTGGAG 2917
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 871 -----Ile-LeuProAlaIleCysLeuLeuThrGlyLysPheIl 883
Db 2918 CTAGGGAATGAGCGGATAACATCTATTCACGCTGTTTGCTTCTCACCGGCAATTCAT 2977
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 883 eMetProGluIleSerAsnLeuAlaSerIleTrp----- 894
Db 2978 CATTCCAACTCTAAGCAACCTTACAAGTGTGTGTAAGTTCGACAAACAGAAAGTGGC 3037
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 895 -----PheIleAlaLeuPheL 900
Db 3038 CGTTTAAGTAGTAGGTTGAGATTCTGTTGGAATGTTCAACACGTTCTTGACACTTTTCC 3097
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 900 euSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGlu- 919
Db 3098 TCTCCATCATTTGCAACTGGAGTGCTTGAACCTTCGATGGAGCGGGTTAGCATCCAAGAAC 3157
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 919 ----- 919
Db 3158 CGTGAAGAGGAGGTAGTAACGTTGACCTCAGCAACTTGAAGTACCTCGCCCAATCG 3217
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 920 -----TrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeu 936
Db 3218 TAGCAAGACTGGTGGCCCAATGAACAATTCGGGTGATCGAGAGTGCTCGGCCCATCTT 3277
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 937 PheAlaValPhe----- 940
Db 3278 TTTGCTGTCTCCGTTCTGACCACCGGTTACTTGTGTAAGACCACTAGCCTCCACAGAG 3337
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 941 -----GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAs 953
Db 3338 GCGGGTAGAAAAACACAGAGAAGGAGGCGCTCCTCAAAAGTCTTAGCTGGAGTAGACACAA 3397
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 953 nPheThrValThrSerLysAlaAsnAspGluGlu-Gly----- 965
Db 3398 CTTCCCGGTAAACAGCAAAAGCAGACGATACTCCCGGAGGAGTTTCAGATCGACCTC 3457
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 966 -----AspPheAlaGluLeuT 971
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 3458 ATCTGTGTTGAAAGTGGCAATGTGTTTTTCGTGCTGTGCTATGAGAATTCGGTGAACTTT 3517
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 971 YrMetPheLysTrpThrThrLeuLeuLeuProProThrThrThrLeuLeuLeu----- 988
Db 3518 ATCTCTCAAAATGACAACTCTCTTAATCCCTCCCACAACTCTGATAATACTGTCTTTAAG 3577
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 988 ----- 988
Db 3578 CCACCTGAAATAGAGAAGTTTACTGTTGAGAGAAATTAGGAGGGTGTTCAGACTATTAT 3637
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 989 ----AsnMetValGlyValValAlaGlyThrSerTyrrAlaIleAsnSerGlyTyrrGlnSer 1007
Db 3638 GACAAACATGTCGGAGTCGTGGCGGAGTTTCAGACGCAATCAACAACGCGCTATGTTCA 3697
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1008 TrpGlyProLeuPhe----- 1012
Db 3698 TGGGGTCCATTTGTTGTACCAGCCTCAGACCCGCGCTCAAAGTCTGCGTTAGTTGTTGCC 3757
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1013 -----GlyLysLeuPhePheAlaPheTrpValIleValHi 1024
Db 3758 GATACCAAGTACCCAGGTAACATCGGCAAACTGTCTTCGCATTCTGGGTCACTCTTCA 3817
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1024 sLeuTyrrProPheLeuLysGlyLeuMetGlyArg----- 1035
Db 3818 TCCTTACCCATCTCTCAAAGTTTGTATGGGAGAGCGGTTTGACAAGAAGCGTAAGACCC 3877
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1036 -----GlnAsnArgThrProT 1041
Db 3878 AGTAAGAAGTAGAAATGGGTAAAGAGTTTCCAAACTACCCCTCACAACACAGGACGCCA 3937
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1041 hrIleValIleValTrpAlaValLeuAlaSerIlePheSerLeuLeuTrp----- 1058
Db 3938 CCATTGTTGCTTGTGTCATCTTTTGCAATCGAATTTCTCACTGTTGGTGGTTTG 3997
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1058 ----- 1058
Db 3998 TCCTCGGGTGTAAACAACAGAACCCAGGTATGAAAACCGTAGCTAAAAGAGTGACCA 4057
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1059 ----ValArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGly 1077
Db 4058 ACGTACGGATGATCCCTTCTTGCCCAACAAACAGGTCCAGTCTCTTAAACATGTGGC 4117
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1078 IleAsnCys 1080
Db 4118 GTGAGTGC 4126
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4118 GTGAGTGC 4126
QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 11
US-09-838-586-2
; Sequence 2, Application US/09838586
; Patent No. 6576818
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/02/US
; CURRENT APPLICATION NUMBER: US/09/838,586
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4612
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-09-838-586-2
Alignment Scores: 9.71e-187 Length: 4612
Pred. No.: 9.71e-187 Length: 4612
```

Score: 1833.00 Matches: 510
Percent Similarity: 42.72% Conservative: 80
Best Local Similarity: 36.93% Mismatches: 78
Query Match: 31.72% Indels: 714
DB: 4 Gaps: 29

US-09-900-237A-30 (1-1080) x US-09-838-586-2 (1-4612)

QY 413 AlaArgLysTyrValProPheValLysLysTyr-Asp----- 424
DB 39 GCGAGGAGATGGGTTCGGTTTTGTGAAGACATTCATCACCTAGGGGGCCGACGCTCTT 98
QY 425 -----IleGluProArgAl 429
DB 99 AAGCGGTGCTGCTCCTCTACCAAGGCAAAACATTCCTCGTTAATGTTGAGCCCGAGGCG 158
QY 429 aProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGln----- 446
DB 159 GCCCGAGTTTATTTCAATGAGAAGATGATTAATTTGAAGGCAAGGTCATATTACAAC 218
QY 446 ----- 446
DB 219 TCGGGTCCGGCGGCTCAAAATAAAGTTACTCTTCTAACTAATAAACTTCCTGTTCCAGG 278
QY 447 --ProSerPheValLysAspArgAlaMetLysArgGluTyrGluGluPheLysIleA 466
DB 279 TACCTAGCTTTGTTAAAGAACGAGAGCCATGAAAGGGAATATGAAGAAATTTAAAGTAA 338
QY 466 rGileAsnAla----- 469
DB 339 GGATCAATGATGGATCGAAACAATTTCTTGCTCTCGGTACTTTTCCCTTATATCTTCT 398
QY 470 -----LeuValSerLysAlaLeuLysValProGluGluGlyTyr 482
DB 399 AAATTTCAATCTAGTTAGTATAGTACGAAAGCTCAGAAACACGAGAGATGG 458
QY 483 IleMetGlnAspGlyThrProTyrProGly-Asn----- 493
DB 459 GTGATGCAAGATGGCACCCATGCGCCGCGAAATCATCGTTTTCGAGTCTCTTTGGTCT 518
QY 494 -----AsnThrArgAspHisPr 499
DB 519 TCTTCTACCCACTAGTTCTACCGTGGGTACCGGGCTTTTAAACACTCGTGATCATCC 578
QY 499 oGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGly----- 516
DB 579 TGGAAATGATTCAAGTCTATCAGGAAGTCCCGGTGCATCGATGTGGATGGCAATTGTGAG 638
QY 516 ----- 516
DB 639 CACTAGTAGACCTTACTAAGTCCAGATAGATCTTTCACGGCCAGCTGAGCTACACCTAC 698
QY 517 --AsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisH 536
DB 699 CGAAAGAGCTGCTCGACTGTCTATGTTTCTCGTGAGAAACGACCTGGTTATCAGCACC 758
QY 536 IsLysLysAla----- 539
DB 759 ATAAGAAAGCCGTTTCTCGACGGAGCTGAACAGATACAAGAGCACTCTTTGCTGGACCA 818
QY 540 -----Gly-AlaMetAsnAlaLeuValArgValSerAlaValLeu 552
DB 819 ATAGTCGTGGTATCTTTTCGGCGGTGCTGAGAAATGCTCTGGTTTCGAGTTTCTCGAGTGT 878
QY 553 ThrAsnGlyGlnTyrMetLeuAsnLeuAspCys----- 563
DB 879 ACTAATGCAACCTTCATATTGAATCTGGATTGCAAGCTCTTACGAGACCAAGCTCAAG 938
QY 564 -----AspHisTyrIleAsnAs 569
DB 939 ACGTCACGAATGATTACGTGGGAAGTATAACTTAGACCTAACTGATCATATACATCAACAA 998
QY 569 nSerLysAlaValArgLysAlaMetCysPheLeuMetAspProAsnLeuGlyPro----- 587

DB 999 TAGCAAGCCCATGAGGAAGGATGTGCTTTTAAATGATCCTCAGTTTGGAACTAGTAA 1058
QY 587 ----- 587
DB 1059 TGTAGTTTATATCGTTCGCTACTCCCTTCGCTACACGAAATAATACCTAGGAGTCAAC 1118
QY 588 ---Gln-ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspA 606
DB 1119 CTAGAAGCTTTGTTATGTTCAATTTCCACAGAGATTTGATGGTATTGATCGTCATGATC 1178
QY 606 rGlyrAlaAsn----- 609
DB 1179 GATAGCTTAATCTTCTTCGAAACATACAAGTTAAAGGTGTCTCTAAACTACCATAACTA 1238
QY 610 ---Arg-AsnThrValPhePheAspIleAsnLeuArgGlyLeu 622
DB 1239 CGAGTACTAGCTATACGATTAGGAAATGTTGTCTTTTGTATATCAATGTTGGATTA 1298
QY 623 AspGlyIleGlnGlyProValTyrValGlyThr----- 633
DB 1299 GATGGACTTCAAGGCCCTGTATATGTAGGCACCTTTACAACAGAGAATACTATAGTTGTA 1358
QY 634 -----GlyCysValPheAsnAr 639
DB 1359 CAACCTAATCTACCTGAAGTTCCGGGACATATACATCCGTGAGGGTGTGTTTCAACAG 1418
QY 639 gThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysPro----- 654
DB 1419 CGAGCATTTGATGGCTACGATCCACCACTCTCTGAGAAACGACCAGAGATGTCACAC 1478
QY 654 ----- 654
DB 1479 AAAAGTTGTCGTCGTAACATACCGATGTAGGTGGTCAGAGACTCTTTGCTGTTTCT 1538
QY 655 -----GlyPheLeuAla---SerLeu-CysGlyGlyLysLysLysAlaAs 668
DB 1539 ACACATGTGATTGCTGGCTTCTTGTGTTGCTGTGTTGCGGAGGTCTCTAGGAAGAA 1598
QY 668 rLysSerLysLysLysArgSerSerLysLysLysSerAsnLysHisVal----- 684
DB 1599 CAAGAAGAAAGTGTACACTACGACCGGAGAACCAACACGACACAA-CCCTTCCAAG 1657
QY 684 ----- 684
DB 1658 ATCCCTTCTTAGTTTCTTCTTTCGTAAGAAAGGGCTTACTCGGAGGTCTTTTATACGG 1717
QY 684 ----- 684
DB 1718 AAAAAAGAAAGATGATGGGCAAAACTATGTCTCTTCTTCGGAATGAGCCTCCAG 1777
QY 685 -----AspSerSerValP 689
DB 1778 AAAATATGCTTTTTTCTTCTTCTACTACCGTTTTTGTATCAGAAAAAGGCTCTGCAC 1837
QY 689 rValPheAsnLeuGluAspIleGluGluGlyValGluGly----- 702
DB 1838 CAGTCTTTGATCTCGAAGAAATCGAAGAGGGCTTGAAGGATACGAAGAATGCTTTTTT 1897
QY 703 -----AlaGlyPheAsp- 706
DB 1898 CCCAGACGTGTCAGAAACTAGAGCTTCTTTAGCTTCTTCCGAACTCTCTATGTTCTT 1957
QY 707 AspGluLysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAla 726
DB 1958 AACGAGAATCGACATTATGTCGAGAGAAATTTTCGAGAACGATTCGACAAATCACCG 2017
QY 727 AlaPheValAlaSer----- 731
DB 2018 GTTTTCAITGCTCTCTTTTAGCTGTAATTACAGCGCTCTTCTTAAAGCTCTTTGCTAAGCC 2077
QY 732 -----ThrLeuMetGluTyrGlyGlyValProGlnSerSe 743

Db 2078 TGTTAGTGGCCAAAGTAACGACAACTTTGATGAAATGTTGGCCTTCTCTGAAGAAC 2137
Qy 743 rThrProGluSerLeuLeuLysGluAlaIleHis----- 754
Db 2138 TAATTCACATCACTGATTAAGAGGCCAATTCAGTTGAACCTATCTTTTACCACCGGAAG 2197
Qy 755 -----ValileSerCysGlyT 760
Db 2198 GACTTCCTTGATTAGGTGTAGTGACTAAATTTCTCCGGTAAGTCGTAAATTAGCTGTGGTT 2257
Qy 760 YrGluAspLysSerGluTrpGlyThrGluIleGlyTrrpIleYrGlySerVal----- 777
Db 2258 ATGAAGAAAAAATCTGAGTGGGGCAAAGAGATCGGATGGATTTATCGGTTCGGTGGCATTA 2317
Qy 777 ----- 777
Db 2318 TCGACACCAATACCTTTTCTTTTGTACTCACCCCGTTTCTCTAGCCTACCTAAATACCCAGC 2377
Qy 778 ---ThrGluAspIleLeuThrGlyPheLysMethHisAlaArgGlyTrrpArgSerValTyr 796
Db 2378 CACACGGAAGATATATTAAACAGGTTTCAAGATGATCTAGAGGGTGGAAATTCGGTTAT 2437
Qy 797 CysMetProLysArg----- 801
Db 2438 TGTGTACCGAAATATGCCTTCTATATATTAATTTGCCAAAGTTCTACGTAAACATCTCCACCTT 2497
Qy 802 -----ProAlaPheLysGlySerAlaProIleAsnLeuSe 813
Db 2498 TAGCCAAATACACATGCTTTTACCGGGCAATCAAAGGGTCCGCTCCAATCAATCTCTC 2557
Qy 813 rAspArgLeuAsnGlnValLeuArgTrrpAlaLeu----- 824
Db 2558 GGATCGGTGACCAAGTTTGTAGATGGGCACTCTGGCCGTAAAGTTTCCAGGGCGAGGTT 2617
Qy 825 -----GlySerValGluIleL 830
Db 2618 AGTTAGAGAGCTAGCCAAACGTGTTCAAAACTCTACCCGTGATGGTTCTGTAGAAATTT 2677
Qy 830 euPheSerArgHisCysProLeuTrpTyrGlyTyrGlyArgLeuLys----- 846
Db 2678 TCCTTAGTCGTCACTGCTCCACTTTGGTATGGTTATGTGGGAAACTGAAATGGACCAAGA 2737
Qy 847 -----P 847
Db 2738 CATCTTTAAAGAAATCAGCAGTACAGGTGAAACCAATACCAATACCACTTTTGACTTT 2797
Qy 847 he-LeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeu 866
Db 2798 ACCCTCGAGAGGCTTGCTTATATCAACACCAATGTTTACCCCTTTCACTCGATCCCTTTA 2857
Qy 867 LeuValTyrCys----- 870
Db 2858 CTCGCCCTATTGTAGAGCTCTCCGAACGAATATAGTTGTGGTAAACAAATGGGAAAGTGGAG 2917
Qy 871 -----Ile-LeuProAlaIleCysLeuLeuThrGlyLysPheI 883
Db 2918 CTAGGGAATGACGGGATAACATCTATTCCAGCTGTTTGTCTCTACCGCAATTCAT 2977
Qy 883 eMetProGluIleSerAsnLeuAlaSerIleTrrp----- 894
Db 2978 CATTCCAACTCTAAGCAACCTTACAAAGTGTGTGATAAGGTTCGACAAACAGAGAGTGGC 3037
Qy 895 -----PheIleAlaLeuPheL 900
Db 3038 CGTTTATAGTAGTAAAGTTGAGATTCGTTGGAATGTTTACACACAGCTTCTTGGCACTTTTCC 3097
Qy 900 euSerIlePheAlaThrGlyIleLeuGluMetArgTrrpSerGlyValGlyIleAspGlu- 919
Db 3098 TCTCCATCATGCAACTGGAGTGTCTGAACCTTCGATGAGCGGGTTAGCATCCCAAGAC 3157
Qy 919 ----- 919
Db 3158 CGTGAAGAGGAGGTAGTAACGTTGACCTCACGAACTTGAAGCTACCTCGCCCCAATCG 3217

Qy 920 -----TrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeu 936
Db 3218 TAGCAAGACTCGTGGCGCAATGAACAATTTCTGGGTGATCGGAGGTGTCTCGCCCATCTT 3277
Qy 937 PheAlaValPhe----- 940
Db 3278 TTTGCTGTCTTCGTTCTTGACACCGCGTACTTGTTTAAGACCCACTAGCCTCCACAGAG 3337
Qy 941 -----GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAs 953
Db 3338 CGCGGTAGAAAAACGACAGAGGAGGCGCTCTCTCAAGTCTCTAGCTGGAGTAGACACCAA 3397
Qy 953 nPheThrValThrSerLysAlaAsnAspGluGlu-Gly----- 965
Db 3398 CTTTACCCTAAACAGCAAAAGCAGACGATACTCTCCGGAGGAGTTTCAGGATCGACCTC 3457
Qy 966 -----AspPheAlaGluLeuT 971
Db 3458 ATCTGTGTTGAAGTGGCATTTGCTGTTTCTGCTGCTGTATGAGAATTCGGTGAACCTT 3517
Qy 971 YrMetPheLysTrpThrThrLeuLeuIleProProThrThrIleLeuIleIle 988
Db 3518 ATCTCTTCAATGACAACTCTCTTAATCCCTCCCAACACTCTGATAACTACTGTCTTAAG 3577
Qy 988 ----- 988
Db 3578 CCATCTGAAATPAGAAAGTTTACCTGTTGAGAGAATTAGGGAGGTTGTGAGACTATTAT 3637
Qy 989 ---AsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSer 1007
Db 3638 GACAAACATGTCGGAGTCTGTCGGCGGAGTTTCAGACGCAATCAACACGGCTAIGGTTCA 3697
Qy 1008 TrpGlyProLeuPhe----- 1012
Db 3698 TGGGTCATGTTTGTGTACCAGCTCAGCACCGGCTCAAAAGTCTGCGTTAGTTGTTGCC 3757
Qy 1013 -----GlyLysLeuPhePheAlaPheTrpValIleValHi 1024
Db 3758 GATACCAAGTACCCAGGTAACATCGGCAAACTGTTCTTCGCACTCTGGGTCACTTCTTCA 3817
Qy 1024 sLeuTyrProPheLeuLysGlyLeuMetGlyArg----- 1035
Db 3818 TCITTTCCCATTCCTCAAGGTTTGTATGGGAGAGCCGTTTGAACAAGCGTGAAGACC 3877
Qy 1036 -----GlnAsnArgThrProT 1041
Db 3878 AGTAAGAAGTAGAAATCGGTAAAGAGTTTCCAAACTACCCCTCACAAAACAGGACGCCA 3937
Qy 1041 hrIleValIleValTrrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrp----- 1058
Db 3938 CCATTGTTGTCTTGTGTCATACATTTTGGATCGATTTTCTCACTGTTGGTGTGTTG 3997
Qy 1058 ----- 1058
Db 3998 TCCTCGCGGTGTAAACAAACAGAAACCCAGGTATGAACCCGTAGCTAAAAGAGTGACCAA 4057
Qy 1059 ---ValArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGly 1077
Db 4058 ACCGTACGGATCGATCCCTTCTTGCCCAACAAACAGGTCAGGTCCTTAACAATGTGGC 4117
Qy 1078 IleAsnCys 1080
Db 4118 GTGGAGTGC 4126

RESULT 12

US-09-221-013A-13
; Sequence 13, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.

```

; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1741
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1741)
US-09-221-013A-13

Alignment Scores:
Pred. No.: 3,11e-179 Length: 1741
Score: 1756.50 Matches: 346
Percent Similarity: 73.18% Conservative: 66
Best Local Similarity: 61.46% Mismatches: 100
Query Match: 30.40% Indels: 51
DB: 4 Gaps: 13

US-09-900-237A-30 (1-1080) x US-09-221-013A-13 (1-1741)
QY 2 AspGlyAspAla---AspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
DB 167 GACGGCGCGCCGCCCGCGCTAAGCCAGGAGAGTGAATGCTCAGTCTGCCAG 226
QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
DB 227 ATTGTGGCGACACTGTTGGCGTCTCGGCCACCGCGACGCTCTTGTTCCTGCAATGAG 286
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
DB 287 TCGCGCTTCCTCCGCTCGCGCCCTGCTACGAGTACGAAACGCAAGGAGGAAACCGAGTGC 346
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
DB 347 TGCCCCAGTGCAGACTAGATACAGAGCACAAGGTGCGCTAGAGTTGAGGCGGAT 406
QY 81 GluGlyAspAspThrAspAlaAspAsp---GlySerAspAsnTyrPro----- 96
DB 407 GAG---GAAGAAGACATGTTGATGACCTGGACATGAATTCATTATAAGCATGGCAAT 463
QY 97 AlaSerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMet 116
DB 464 GGCAAGGTCCAGAG-----TGGCAGATA 487
QY 117 AsnThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeu 136
DB 488 CAGAGACAGGGGAA-----GATGTTACCTGCTTCATCT 523
QY 137 SerLysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln 156
DB 524 TCTCGCCACCAACATCCGATTCGCCGT-----CTGACAAGTGGGCAACAG 571
QY 157 MetSerGlyGluIleProGlyAlaSerProCysPheHisHisMetMetSerProThrGlyAsn 176
DB 572 ATCTCAGGAGATCCCTGATGCTCCCGCCGATTCGCCATTCCTATC----- 616
QY 177 IleSerArgArgAlaProPheProTyrValAsnHisSer----- 189
DB 617 -----CGCAGCGGACATCAAGCTATGTTGATCCCAAGTCTCCAGTCTCTGAGGATT 670
QY 190 ProAsnProSerArgGluPhe---SerGlySerIleGlyAsnValAlaTrpLysGluArg 208
DB 671 GTGACCCCTCCAAAGGACTTGAATTCCTATGGGATTAAACAGTGTGACTGGCAAGAAAGA 730

```

```

QY 209 ValAspGlyTyrLysMetLysGlnAspLysGlyAlaIleProMetThrAsnGlyThrSer 228
DB 731 GTTGCCAGCTGGAGAACAGCAGGACAAAATAATATGTCAGGTAGCTAAT----- 781
QY 229 IleAlaProSerGluGlyArgAlaIleThrAspIleAspAlaSerThrGluTyrAsnMet 248
DB 782 ---AAATATCCAGAGGCAAGGGGGA---GACATGGAA---GGGACTGGTTCAATGCT 832
QY 249 GluAspAlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLysValProIleAla 268
DB 833 GAAGATATCCAAATGCTGATGATGACGCTCTACCTCTGAGCGCATAGTGCCTATCCCT 892
QY 269 SerSerLysIleAsnProTyrArgMetValIleValLeuArgLeuValValLeuSerIle 288
DB 893 TCAAAACCACTCAACCTTTACCGAATGTTATCATCTCCGCTTATCATCTCATGATGTC 952
QY 289 PheLeuHisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeuTyrPheLeuSer 308
DB 953 TTCTTCCAATATCGTGTCACTCATCCAGTGGGATGCTTATGGATTGTGGCTAGTATCT 1012
QY 309 ValIleCysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPheProLysTrpPhe 328
DB 1013 GTTATCTGTGAATTTGGTTGGCTTATCTCGCTCCTAGATCAATTCCAAAGTGGTAC 1072
QY 329 ProIleAsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGlu 348
DB 1073 CCGATAAACCGTGAACATACCTTGACAGGCTTGATTTAGATATGATAGGAGGGAGAG 1132
QY 349 ProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspProLeuLysGluPro 368
DB 1133 CCATCACAGCTTGCTCCCATTCATGCTTTTGTGAGTGGTGGATCCACTAAAGGAACCT 1192
QY 369 ProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLys 388
DB 1193 CCTCTGATCACAGCAACACACTGTTTGTCCATTCTGGCTGGATTAACCTGTGACAAA 1252
QY 389 ValSerCysTyrValSerAspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGlu 408
DB 1253 GGTGATGCTATGTTCTGACCATGGTTCAGCTATGTTAACTTTTGAGGCTCTGTGAGAA 1312
QY 409 ThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAspIleGluProArg 428
DB 1313 ACTGCAGAAATTTGCTAGGAAGTGGGTTCCGTTTTCAGAAAGCACAAATATTGAACCCGA 1372
QY 429 AlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGlnProSer 448
DB 1373 GCTCCAGAGTTTACTTTGCTCAAAAAATAGATTACCTGAAGAGGGGTGGACCATGCTGTC 1432
QY 449 PheValLysArgArgAlaMetLysArgGluTyrGluGluPheLysIleArgIleAsn 468
DB 1433 TTTGTTAAGAAAGCGGCAATGAGAGAGATGAGAGATTAAGAAATTCAGGTACGGATCAAT 1492
QY 469 AlaLeuValSerLysAlaLeuLysValProGluGluGlyTyrIleMetGlnAspGlyThr 488
DB 1493 GCTCTTTGTCGAAGGCACAAAAAGTACCTGAAGAGGGGTGGACCATGCTGTCGTCACCT 1552
QY 489 ProTyrProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis 508
DB 1553 GGTGTGCTCGGGAATAACCAAGGGATACCTCGGCATGATTCAAGGTGTTCTTGGGGCAC 1612
QY 509 SerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu 528
DB 1613 AGTGGTGGGCTTGACACTGATGGTAACGAGTTGCCACGGCTGCTACGCTCTCTCGTGA 1672
QY 529 LysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAlaLeuValArgVal 548
DB 1673 AAGAGCCAGGATTCAGCATCACAAAGAGGTGTCGAATGAATGATTCATTGATTCGTGA 1732
QY 549 SerAlaVal 551
DB 1733 TCTGCTGTG 1741

```

```
RESULT 13
US-09-313-294A-6485
; Sequence 6485, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6485
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700351882H1
US-09-313-294A-6485
Alignment Scores:
Pred. No.: 1.6e-42 Length: 281
Score: 478.00 Matches: 88
Percent Similarity: 98.92% Conservative: 4
Best Local Similarity: 94.62% Mismatches: 1
Query Match: 8.27% Indels: 0
DB: 4 Gaps: 0
US-09-900-237A-30 (1-1080) x US-09-313-294A-6485 (1-281)
QY 854 IleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuPro 873
Db 2 ATCAACACCACTCTACCGCGTCACGTCCTCCGCTCCTCATTTACTGTATCTGCT 61
QY 874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluLeuSerAsnLeuAlaSerIle 893
Db 62 GCACATCTGCTGTCACGGGAGATTTCATATCCAGAGATCAGCAACTTCGCTAGTATC 121
QY 894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSer 913
Db 122 TGGTTCATCTCTCTTCATCTCGATCTTCGCCACGGGTATCTCTGGAGATGAGTGGAGC 181
QY 914 GlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTyrValIleGlyIleSer 933
Db 182 GCGTGGGCATCGACGAGTGTGGAGGAGCAGCAGTCTCTGGTTCATCGGAGGCATCTCC 241
QY 934 AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysVal 946
Db 242 GCCACCTCTTCGCGCTTCACGGGCTCTCTCAAGGTG 280
RESULT 14
US-09-313-294A-4753
; Sequence 4753, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4753
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350285H1
US-09-313-294A-4753
Alignment Scores:
Pred. No.: 1.32e-38 Length: 291
Score: 442.00 Matches: 81
Percent Similarity: 89.69% Conservative: 6
Best Local Similarity: 83.51% Mismatches: 9
Query Match: 7.65% Indels: 1
DB: 4 Gaps: 0
US-09-900-237A-30 (1-1080) x US-09-313-294A-4753 (1-291)
QY 790 ArgGlyTrpArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAlaPro 809
```

```
; OTHER INFORMATION: Incyte ID No. 6476212 700348982H1
; NAME/KEY: unsure
; LOCATION: 64, 189
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4753
Alignment Scores:
Pred. No.: 2.3e-42 Length: 300
Score: 477.00 Matches: 92
Percent Similarity: 97.00% Conservative: 5
Best Local Similarity: 92.00% Mismatches: 3
Query Match: 8.26% Indels: 1
DB: 4 Gaps: 0
US-09-900-237A-30 (1-1080) x US-09-313-294A-4753 (1-300)
QY 886 GluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeuSerIlePheAlaThr 905
Db 2 CAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTCTTCATCTCGATCTTCGCCACG 61
QY 906 GlyIleLeuGluMetArgTyrTrpSerGlyValGlyIleAspGluTyrTrpArgAsnGluGln 925
Db 62 GGNATCCTGGAGATGAGTGTGGAGCGGGTGGGCATCGACGAGTGTGGAGGAGCAGCAG 121
QY 926 PheTyrValIleGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLys 945
Db 122 TTCGGGTGATTCGGGGCATCTCCGCGCATCTTCGCGGTTCAGGGCTCTCTCAAG 181
QY 946 ValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGly 965
Db 182 GTGCTGGCGCATCGACACCACTTCACCGTCACCTCCAGGCTCGGACGAGGAGCGGC 241
QY 966 AspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeuLeuProThrThrIle 985
Db 242 GACTTCGCGGA-CTGTACATGTTCAAGTGGACGACGCTCTTATCCGCCGCCACCATC 300
RESULT 15
US-09-313-294A-5519
; Sequence 5519, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5519
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350285H1
US-09-313-294A-5519
Alignment Scores:
Pred. No.: 1.32e-38 Length: 291
Score: 442.00 Matches: 81
Percent Similarity: 89.69% Conservative: 6
Best Local Similarity: 83.51% Mismatches: 9
Query Match: 7.65% Indels: 1
DB: 4 Gaps: 0
US-09-900-237A-30 (1-1080) x US-09-313-294A-5519 (1-291)
QY 790 ArgGlyTrpArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAlaPro 809
```

D _b	3	CGANGCTGGCATCAATCTACTGCAGTGCTTAAGCAGCAGCTTTTCAAGGGANCTGCTCCT	62
Q _y	810	- IleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluI	829
D _b	63	ATNCNACCTTTTCGGATCGTTTGAATCAAGTCTTCGGTGGGCTCTTGGTTCATTGSAAT	122
Q _y	829	eLeuPheSerArgHisCysProLeuTrpTyrGlyTyrGlyGlyArgLeuLysPheLeuG1	849
D _b	123	TCITTTTCAGCAGGCATTGTCCTCCATATGGTATGGCTATGGAGTTCGGCTTAATTCCTGGA	182
Q _y	849	uArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuValTy	869
D _b	183	GAGATTTTCGCTATATCANCACCAACAATTTATCCACTCANATCAATTCGNGCTCCTCNTGTGA	242
Q _y	869	rCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetPro	885
D _b	243	CTGCATATGGCAGCAGATNTGTTCTCTCACTGGGAAGTTTCATCATCCCA	291

Search completed: August 22, 2004, 22:25:26
Job time : 300 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 22, 2004, 20:43:24 ; Search time 1097 Seconds
(without alignments)
4836.011 Million cell updates/sec

Title:
Perfect score: 5778
Sequence: 1 MGDADALSKSRHGAGDVQ.....VDFPTRLAGPNLTGGINC 1080

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3228839 seqs, 245606551 residues
Total number of hits satisfying chosen parameters: 6457678

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0900237/runat_18082004_081516_14735/app_query.fasta_1.1223
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl6sum62
-TRANS-human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZB=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0900237@cgn2_1 1 750 @runat_18082004_081516_14735
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=130 -WARN TIMEOUT=130 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:**
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:**
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:**
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:**
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:**
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:**
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:**
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:**
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:**
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:**
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:**
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:**
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:**
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:**
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:**
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:**
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:**
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:**
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:**
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
------------	-------	-------------	-----------	----	-------------

ALIGNMENTS

RESULT 1
US-09-900-237-29
; Sequence 29, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 3626
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-900-237-29

1	5778	100.0	3626	9	US-09-900-237-29	Sequence 29, Appl
2	5452.5	94.4	3962	17	US-10-437-963-72402	Sequence 72402, A
3	5423.5	93.9	3799	13	US-10-627-132-21	Sequence 21, Appl
4	5423.5	93.9	3799	15	US-10-267-459-5	Sequence 5, Appl
5	5423.5	93.9	3799	15	US-10-209-059-21	Sequence 21, Appl
6	5421.5	93.8	3746	15	US-10-160-719-29	Sequence 29, Appl
7	5421.5	93.8	3746	15	US-10-160-719-49	Sequence 49, Appl
8	5421.5	93.8	3773	13	US-10-627-132-9	Sequence 9, Appl
9	5421.5	93.8	3773	15	US-10-209-059-9	Sequence 9, Appl
10	5421.5	93.8	3773	15	US-10-160-719-5	Sequence 5, Appl
11	5149	89.1	3704	13	US-10-627-132-13	Sequence 13, Appl
12	5149	89.1	3704	15	US-10-209-059-13	Sequence 13, Appl
13	5149	89.1	3704	15	US-10-160-719-57	Sequence 57, Appl
14	5129	88.8	3534	13	US-10-425-114-2553	Sequence 2553, Ap
15	5107.5	88.4	4284	17	US-10-437-963-39762	Sequence 39762, A
16	4499	77.9	3614	15	US-10-229-193-9	Sequence 9, Appl
17	4029	69.7	3723	13	US-10-424-599-106896	Sequence 106896,
18	4021.5	69.6	3563	9	US-09-900-237-25	Sequence 25, Appl
19	3993.5	69.1	3725	13	US-10-627-132-45	Sequence 45, Appl
20	3993.5	69.1	3725	15	US-10-209-059-45	Sequence 45, Appl
21	3993.5	69.1	3725	15	US-10-160-719-13	Sequence 13, Appl
22	3993.5	69.1	3725	15	US-10-160-719-21	Sequence 21, Appl
23	3993.5	69.1	3725	15	US-10-160-719-41	Sequence 41, Appl
24	3988	69.0	3753	15	US-10-160-719-33	Sequence 33, Appl
25	3988	69.0	3753	15	US-10-160-719-53	Sequence 53, Appl
26	3988	69.0	3780	13	US-10-627-132-1	Sequence 1, Appl
27	3988	69.0	3780	15	US-10-209-059-1	Sequence 1, Appl
28	3988	69.0	3780	15	US-10-160-719-9	Sequence 9, Appl
29	3963.5	68.6	3603	15	US-10-229-193-5	Sequence 5, Appl
30	3959.5	68.3	3673	15	US-10-229-193-11	Sequence 11, Appl
31	3946	68.5	3851	16	US-10-393-840-69	Sequence 69, Appl
32	3890	67.3	3443	13	US-10-627-132-29	Sequence 29, Appl
33	3870.5	67.0	3448	17	US-10-437-963-15093	Sequence 15093, A
34	3861	66.8	3589	13	US-10-425-114-24930	Sequence 24930, A
35	3844.5	66.5	3786	9	US-09-900-237-7	Sequence 7, Appl
36	3844.5	66.5	3813	13	US-10-627-132-17	Sequence 17, Appl
37	3844.5	66.5	3813	15	US-10-209-059-17	Sequence 25, Appl
38	3844.5	66.5	3813	15	US-10-160-719-25	Sequence 45, Appl
39	3844.5	66.5	3813	15	US-10-160-719-45	Sequence 45, Appl
40	3820	66.1	3568	13	US-10-627-132-41	Sequence 41, Appl
41	3820	66.1	3568	15	US-10-209-059-41	Sequence 41, Appl
42	3820	66.1	3568	15	US-10-160-719-1	Sequence 1, Appl
43	3818	66.1	3776	9	US-09-900-237-3	Sequence 3, Appl
44	3818	66.1	3936	9	US-09-900-237-9	Sequence 9, Appl
45	3818	66.1	3969	13	US-10-627-132-49	Sequence 49, Appl

Alignment Scores:

```

Pred. No.:      0      Length:      3626
Score:      5778.00      Matches:      1080
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      9      Gaps:      0

US-09-900-237A-30 (1-1080) x US-09-900-237-29 (1-3626)

QY      1      MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln      20
DB      60      ATGGACGGCGACGCGACGCCCTGAAGTCCGGAGGACACGGGGCGGGGACGCTGTGCCAG      119

QY      21      IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal      40
DB      120      ATCTCGCGCGACGGGCTGGGACCCACGTTGGACGGCGAGCTCTTACCGCCCTGCCACGTC      179

QY      41      CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla      60
DB      180      TGCCGCTTCCCGTCTCGCGCCCTCTGTACGAGCAGCAGCGCAAGGGGCGACCCAGGCC      239

QY      61      CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu      80
DB      240      TGCCCTCCAGTGCACAGCCAAAGTACAAAGCGCCACAGAGGCGAGCCAGCGATCCGCGGGAG      299

QY      81      GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr      100
DB      300      GAAGGCGGACGACACTGATGCCGATGATGTAGTACTTCAACTACCCCTGCATCTGGCACT      359

QY      101      GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly      120
DB      360      GAGGACCAAGACAGAGATTGCTGACAGGATGCCAGCTGGCGCATGACACACCGGGGGC      419

QY      121      SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp      140
DB      420      AGTGGCAATTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCTCTCCAAAGTATGAC      479

QY      141      SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu      160
DB      480      AGTGAGAGATCCCTTAGGGGATACCTCCCTTCCAGTCCACCAAGCAGCAGATGTCAGAGAA      539

QY      161      IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg      180
DB      540      ATCCCTGGAGCTTCGCTGATCATCACATGATGTCCTCCCTACGGGGAACATCAGCAGAGCT      599

QY      181      AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle      200
DB      600      GCTCCGTTTCCCTATGTGAATCATTCACCAATCCGTCGAAGGAGTTCCTCCGGCAGTATT      659

QY      201      GlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAla      220
DB      660      GGGAAATGTTGCTGGAAAGAGAGATTGATGGCTGGAAATGAACGAGACAAAGGGTGG      719

QY      221      IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle      240
DB      720      ATTCCCATGACTAATGGACAAAGCATTCCTCTCTGAAGGTCCGGCGAGCTACTGACATC      779

QY      241      AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro      260
DB      780      GATGCATCTACTGAATACAACATGGAAGACGCTTTTACTGAATGATGAACATCCGCCGCT      839

QY      261      LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal      280
DB      840      CTATCTAGAAAGTCCCATTTGCTTCTCCAAAATPAATTCCTCAGAAATGTCTATTGTT      899

QY      281      LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn      300
DB      900      CTGCGTTTGGTTGTTCTAAGCATCTTCTCGACTACCGCTCTCACAAATCTCTGTGCTAAT      959

QY      301      AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle      320

```

```

DB      960      GCATACCCACTGTGGCTTTTATCTGTTATATGAGATTGGTTTGTCTTTATCTCGGATA      1019

QY      321      LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrIleuAspArgLeuAla      340
DB      1020      CTGGATCAGTTTCCGAAGTGGTTTCCAATCAACCGGAGACCTACTCTTGATAGATGGCT      1079

QY      341      LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer      360
DB      1080      TTAAGGTATGACCGAGAGGTGAACGCTCTCAGTTGGCTGCTGTGACATATTGTCACT      1139

QY      361      ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu      380
DB      1140      ACAGTCGACCCCTTGAAGGAGCCACTATCGTCACTGCGCAACACTGTGTCATCCATCTT      1199

QY      381      AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyIleSerMet      400
DB      1200      GCTGTGTATTATPCCGTGGACAAAGTCTCTGCTATGATCTCTGATGACGAGCTTCAATG      1259

QY      401      LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal      420
DB      1260      CTGACTTTTGGAGCATTTGGCTGAGACTTCAGAGTTTGTAGGAATGGGTACCATTTGTG      1319

QY      421      LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr      440
DB      1320      AAGAAGTATGACATTTGAACCCAGAGCTCCGAGTTTCTTTTGGCAGAAAAATTCATTAC      1379

QY      441      LeuLysAspLysValGlnProSerPheValLysAspArgArgAlaMetLysArgGluTyr      460
DB      1380      CTGAAGACAAAGTCCAGCTTCATTTGTTAAAGACCCCGGGCCCATGAAGAGAGAAATAT      1439

QY      461      GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu      480
DB      1440      GAAGAATTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCATTTGAAGTCCCCGAGAA      1499

QY      481      GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly      500
DB      1500      GGATGGATCATGCAAGATGGCACACCATGGCCAGGAAACAATACCAGGGATCATCTGGA      1559

QY      501      MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro      520
DB      1560      ATGATTCAGGTTTCCCTTGGTCACAGTGGTGGCTGTGATCTAGGGTAAATGAGTCCCTC      1619

QY      521      ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly      540
DB      1620      CGTTTATTTATGTCTCGTGAAAGCGTCTCGGTTTCCAGCACCAAGAGGCTGTG      1679

QY      541      AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn      560
DB      1680      GCCATGAATGCCCTTGTCTCGTCTCTCAGCTGTCTTACTAATGGACAATACATGTTGAAT      1739

QY      561      LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu      580
DB      1740      CTTGATTGTGATCACTACATCAACACAGCAGGCTGTCCGAGAGAGCTATGTGCTTCCTA      1799

QY      581      MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly      600
DB      1800      ATGGATCCAAACCTAGGTCCGCAAGTCTGTTATGTGCAAGTTCCTCCACAAAGGTTTGA      1859

QY      601      IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg      620
DB      1860      ATTGATGAATATGATCGATATGCAACAGGAAACACTGTCTTTTGTATTAATTAACCTGAGG      1919

QY      621      GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr      640
DB      1920      GGCCTTGACGCAATTCAGGACCGATTATGTGGGAATCTGGTTGTGTTTTCAACAGAACTA      1979

QY      641      AlaIleTyrGlyTyrGluProIleLysAlaLysLysProGlyPheLeuAlaSerLeu      660
DB      1980      GCTATCTATGTTATGAGCCCCCAATTAAGCGGAAGAACCCAGGTTTCTTGGCATCACTA      2039

QY      661      CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer      680
DB      2040      TGTGGGGGCAAGAGAGGCAAGCAAGTCAAGAAAGAGGAGCTCAGATAAGAAAAAGTCG      2099

```

```
QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db 2100 AACAGCATGTGGACAGTCTGTGTCCAGTATTCAATCTCGAAGACATAGAGGAGGGTGT 2159
QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db 2160 GAAGGTGCTGGTTTGNATGATGAGAATCAGTCTCATGTCTCAATGAGCTTAGAGAAG 2219
QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
Db 2220 AGATTTGGCCAGTCAGCAGCATTTGTTCCTCCACTCTGATGGAATATGGTGTTCCT 2279
QY 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db 2280 CAGTCTCCACTCCAGAAATCTCTTTGAAGAAGCTATCCATGTCATAAGTTGTGGCTAT 2339
QY 761 GluAspLysSerGluTrpGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
Db 2340 GAGGACAAAGTCTGAATGGGAACTGAGATTGGTTGGATCTATGGATCTGTACAGAGAT 2399
QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetProLys 800
Db 2400 ATTCTTACTGGATTCAAGATGCACGCAAGAGGTGGCGTTTCAGTCTATTGCATGCCCAAG 2459
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db 2460 CGCCCAAGCTTCAAGGATCTGCCCATCAATCTTTCAGATCTGTGAACCAAGTCTG 2519
QY 821 ArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyrGly 840
Db 2520 CGGTGGGTCTCGGTCTCTGTGAAATCTTTTCAGCCGGCATTCGCCCTTATGGTATGGC 2579
QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db 2580 TACGGAGGGCGCCTCAAGTTCTCTGGAGAGATTGCTTTACATCAACACCACCATTTACCCA 2639
QY 861 LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2640 CTAACTCTCTCCCGCTTCTAGTCTATGTATATTGCTCTATCTGTCTGTCTCAGTGGGA 2699
QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeu 900
Db 2700 AAGTTCAATATCCAGAGATTAGCAACTTCGCCAGTATCTGTTTCATTTGGCTTTCCTT 2759
QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrp 920
Db 2760 TCAATTTTCGCCACTGTGTATCCTTGAGATGAGGTGGAGTGTGTGGCAATTCAGAGTGG 2819
QY 921 TrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaValPhe 940
Db 2820 TGGAGGAATGAACAGTTCTGGGTCAITGGAGGTATCTCTGCACATCTGTGTGGCGTCTTT 2879
QY 941 GlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db 2880 CAGGTCTCTGAAGTGTCTTGGCGTATCGACACCAACTTCATCTGTCACTCAAGGCT 2939
QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
Db 2940 AATGACCAAGAAAGCGACTTTGTCTGAGCTCTACATGTTCAAAGTGGACGAGCTTCTCATC 2999
QY 981 ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db 3000 CCTCCGACGACCAATTTTGAATCAATTAATGATGGTGGTGTGGTGTGGCAGCTCTCAGCC 3059
QY 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
Db 3060 ATCAACAGTGGTACCAATCATGGGGCGCTCTTTGGGAGCTCTTCTTTGGCTTCTGG 3119
QY 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3120 GTGATTGTCTACTTATACCATTCCTCAAGGGTCTTATGGGCGAGCAAAACCGCACACCG 3179
```

```
QY 1041 ThrIleValIleValTrpAlaValLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
Db 3180 ACATTTGTCTATCTCTGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3239
QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3240 GTTGATCCATTCATCTACCGGTCTCGCTGGCCCAATATCCAAACCTGTGGCATCAACATGC 3299

RESULT 2
US-10-437-963-72402
; Sequence 72402, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 72402
; LENGTH: 3962
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72784C.1
US-10-437-963-72402

Alignment Scores:
Pred. No.: 0 Length: 3962
Score: 5452.50 Matches: 999
Percent Similarity: 97.22% Conservative: 52
Best Local Similarity: 92.41% Mismatches: 29
Query Match: 94.37% Indels: 1
DB: 17 Gaps: 1

US-09-900-237A-30 (1-1080) x US-10-437-963-72402 (1-3962)
QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 373 ATGACGGCCACCGGATGCGTGAAGTCGGGAGGACCGGAGCGGCGAGCGCGTGCAG 432
QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 433 ATCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 492
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 493 TGGCGCTTCCGGTGTGCGCCCTGCTACAGTACGAGCGCAAGGATGCGACCCAGGCT 552
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 553 TGCCCCCAGTGCAGAACCAAGTACAAAGCGCCAAAGGGAGCGCGCGATCCCGTGGGGAG 612
QY 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 613 GAAGGGCGAGATGATGATGCTGATGATGCTGATGATGCTGATGATGCTGATGATGCTGATG 672
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
Db 673 GCGGACCAAGAGCAGAGAGATTGCTGATGATGATGATGATGATGATGATGATGATGATG 732
QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 733 GGTGGAGACGTGGCGCGCTCCCAAGTATGACAGTGGCGGAGATCGGGCTCCACCAAGATGAC 792
```

Qy 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
 Db AGTGGCGAGATCCCTCGGGATACATCCTTCAGTCACTAATAGCCAGATCTCGGAGAA 852
 Qy 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
 Db ATCCCTGAGGCTTCCCTTCATCATATGATGTCCTTACCGGAAACATTTGGCAAGCGT 912
 Qy 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
 Db GCTCATTTCCCTATGTGAACCAITTCACCAATCCATCAAGGAGTTCTCTGGCAGCAAT 972
 Qy 201 GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyAla 220
 Db GGAAATGTGCTCGAAAGAAAGAGTGTGATGCTGGAACCTGAACGAGGACAAAGGAGCA 1032
 Qy 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
 Db ATTCCCATGACCAACGGGACCAAGCATTCGCCCTTCGAAAGTTCGGGAGTTGGTGATATC 1092
 Qy 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
 Db GATGATCCATCATATCAATATGAAGATGCTTACTGAATGATGAACCTCGCCAGCCT 1152
 Qy 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
 Db CTCTCTGAAAGTTCCCTTCATCCAGAAATAAATCCCTACAGAAATGCTCATTTGTT 1212
 Qy 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
 Db CTGGTGTGGTGTCTAAGCATTTTCTTCACTACCGTATATTACGAATCCTGTGGCAAT 1272
 Qy 301 AlaTyrProLeuTrpLeuSerValIleCysGluIleTyrPheAlaLeuSerTrpIle 320
 Db GCGTATCCGCTCTGGCTTTATCTGTTATATGTGAGATTGGTTGCTTTGTCTGGATA 1332
 Qy 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
 Db TTGGATCAGTTCCCGAAGTGGTTTCCAAATCAACCGTGAACCTACTTGATAGGCTGGCA 1392
 Qy 341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
 Db TTGAGGTATGACAGAGAAGGTGAGCCATCTCAGTTGGCTGCTGTTTGACATTTTGTGCA 1452
 Qy 361 ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu 380
 Db ACAGTCGACCCATGACGAACCTCTCTTGTACTGCGCAATACCGTCTGTCCATTTCTT 1512
 Qy 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
 Db GCTGTTGATTACCCAGTGGACAAAGGCTCTTGTCTATGTTATCTGACGATGGTGTGCAATG 1572
 Qy 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
 Db CTGACTTTTGATGCAATGGCTGAGACTTCAGAGTTGTAGAAAGTGGGTTCCTATTCGTT 1632
 Qy 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
 Db AAGAAGTATACATGACCAAGAGCTCTCTGAGTGGTACTTCTCCCAAGAAATTCGATTAC 1692
 Qy 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr 460
 Db TTGAAGACAAAGTCCACCCCTTCATTTGTTAAAGACCGCTCGTGCATGAAGAGAAATAT 1752
 Qy 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
 Db GAAAGATTCAAAGTTAGGATAAATGGCCTTGTGTCTAAGGCACAGAAAGTCCCGAGGAG 1812
 Qy 481 GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 500
 Db GGATGGATCATGCAAGATGGCACACCATGGCCAGGAAACAATACTAGGGACCATCCTGGA 1872
 Qy 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520

Db 1873 ATGATTCAAGTTCCTTCCTGGTCCAGTGGTGGCTTTGATACCGAGGTAATGAGCTTCC 1932
 Qy 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 540
 Db CCGCTGGTCTACGTATCTCGTGAAGGCTCCTGGGTTTCAGACCAACAAGAGGCTGGT 1992
 Qy 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
 Db GCCATGAATGCTCTTGTTCGTCTCAGCTGTGCTTACCAATGGACAGTACATGTTGAAT 2052
 Qy 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
 Db CTTGATTTGATCACTACATCAACAACAGCAAGGCTCTCGCGGAAGCTATGTGCTTCCT 2112
 Qy 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
 Db ATGGATCCAAACCTAGGAAGGAGTGTCTGTATGTTCAGTTCCCAACAAAGGTTTCGATGT 2172
 Qy 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
 Db ATTGATAGCAATGATCGATATCGNACAGGAACACTGTGTTTTCGATATTAACTTGAGG 2232
 Qy 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
 Db GGTCTTGATGCATCCAAAGGACCAAGTTTATGTGGAACTGGTTGTGTATTCACACAGAACT 2292
 Qy 641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGly---PheLeuAlaSer 659
 Db GCTCTATATGTTATGAACCCCAATTAAGCAGAAAGAAAGGGAAGTTTCTTGTTCATCA 2352
 Qy 660 LeuCysGlyLysLysLysLysAlaSerLysSerLysLysArgSerAspLysLysLys 679
 Db CTATGTGGGCGCAGGAGAGCAAGCAAGTCAAGAAGAAAGAGTTCGACACAGAAAGAG 2412
 Qy 680 SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGly 699
 Db TCAAAACAGCACGTGGACAGTGTGTCAGTTCCTTCAATCTTGAAGATATAGAGAGGCT 2472
 Qy 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
 Db GTTGAAGTGTCTGATTCGATGATGAGAAATCATCTTTATGTCTCAAAATGAGCTTGGAG 2532
 Qy 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyVal 739
 Db AAGAGATTGGCCAGTCTGACAGCGTTGTTGCCCTCCACTCTCATGGAATATGGTGGTGT 2592
 Qy 740 ProGlnSerSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGly 759
 Db CCTCAATCTGCAACCCAGAAATCTCTTTGAAAGAAAGCTATCCATGTGATAAGTTGTGGC 2652
 Qy 760 TyrGluAspLysSerSerGluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGlu 779
 Db TATGAGGACAAAGACCGAATGGGAACTGAGANTGGGTGGATCTACGGTTCCTGTCAGAAA 2712
 Qy 780 AspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetPro 799
 Db GATATTCCTCACTGGATTCAGATGCATGCGGAGGCTGGAGATCAATCTACTGTCATGCC 2772
 Qy 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
 Db AAGGCCCCAGCTTTCAAGGGGTCTGCTCTATCAATCTTTTCAGATCGCTTTAAACCAAGTG 2832
 Qy 820 LeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyr 839
 Db CTTGCTGGGCACTTGGTTCTGTGAAATCTTTTCAGTCGCCATGTGCCCATATGGTAC 2892
 Qy 840 GlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
 Db GGCTATGGAGGACGCTTAAGTTCCTGGAGAGATTTGCTTACATCAACACACCATATTAT 2952
 Qy 860 ProLeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879

```
Db 2953 CCATTGACATCGATCCCGCTTCTCATATACTGTGTGTTTGCCTGCTATCTGTGTTGCTCACT 3012
Qy 880 GtlylvsPheIleMetProGluIleSerAsnLeuAlaSerIleTrrPheIleAlaLeuPhe 899
Db 3013 GGGAGGTTTCATCCCGAGATTAGCAACTTTCCTAGTATTGGTTTCATCTCTCTCTTC 3072
Qy 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrrPserGlyValGlyIleAspGlu 919
Db 3073 ATTTCATTTTGGCCTGGTATCTTGAGATGAGTGGATGGTGGTGGATCGATGAG 3132
Qy 920 TrrPrrArgAsnGluInPheTrrPrrValIleGlyIleSerAlaHisLeuPheAlaVal 939
Db 3133 TGGTGGAGGAATGAACAGTCTGGGTATTGAGGATATCTGGCATCTTTTGGCGTC 3192
Qy 940 PheGlnGlyLeuLeuLysValIleAlaGlyIleAspThrAsnPheThrValThrSerIys 959
Db 3193 TTCAGGGTCTCCTCAAGGTGCTTGCTGGTATCGACCAATTTCACTGTCACTCAAG 3252
Qy 960 AlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrrPrrThrLeuLeu 979
Db 3253 GCTTCTGATGAAGTGGGACTTTGCTGAGCTCTACATGTTCAAGTGGACACACGCTTCTC 3312
Qy 980 IleProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyr 999
Db 3313 ATCCACCGCAGACCACTCTTGATCATTAACCTGGTTCGGTGTGTTGCTGTATCTCATAC 3372
Qy 1000 AlaIleAsnSerGlyTrrGlnSerTrrGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3373 GCGATCAACAGCGGCTACCAATCATAGGGACCGCTCTTTGGCAAGCTCTTTTGGCCCTC 3432
Qy 1020 TrrPrrIleValIleValTrrPrrAlaValLeuLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
Db 3433 TGGGTGATGTCACCTTGATCCCTCTCTCAAGGTCTTATGGGTTCGGCAAAACCGCACT 3492
Qy 1040 ProThrIleValIleValTrrPrrAlaValLeuLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1059
Db 3493 CCGACCATCGTGTGTTTGGGCAATCTTCTGGCTTCGATCTTCTCAATTCGTTGGGT 3552
Qy 1060 ArgValAspProPheThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
Db 3553 CGCATCATCATTCACACCGGTGTACCGGCCGAGATACCCAAACATGTGGCATCAAC 3612
Qy 1080 Cys 1080
Db 3613 TGC 3615

RESULT 3
US-10-627-132-21
; Sequence 21, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; PRIOR FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 3799
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: 3757_3775, 3777, 3782
; OTHER INFORMATION: n = A,T,C or G
US-10-627-132-21

Alignment Scores:
Pred. No.: 0 Length: 3799
Score: 5423.50 Matches: 996
Percent Similarity: 96.48% Conservative: 46
Best Local Similarity: 92.22% Mismatches: 37
Query Match: 93.86% Indels: 1
DB: 13 Gaps: 1

US-09-900-237A-30 (1-1080) x US-10-627-132-21 (1-3799)
Qy 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 238 ATGAGGCGCGACGCGGACCGCGTGAAGTCCGGGAGGCGCGGGGAGGCGAGGTGTGCCAG 297
Qy 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 298 ATCTGCGCGGATGGCGTGGGCACTACGCGGAGGAGGAGACGCTTCCCGCTGGACGTC 357
Qy 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 358 TGGCGGTTCGCGGTGTCGCGCCCTGCTACGAGTACGAGCGCAAGGAGGCGACACGCG 417
Qy 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 418 TGCCCCCAGTGTCAAAAACAGTCAAGCGCCACAGGCGAGTCCAGCGATCCGAGGGAG 477
Qy 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerClyThr 100
Db 478 GAAGGAGACGATCTGATGCGGATGCTAGCGACTTCACTACCTCGCTGCTGCGCAAT 537
Qy 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrrArgMetAsnThrGlyGly 120
Db 538 GACGACCAAGACGAGAGATGCTGACAGGATGCGAGCTGGCGCATGAATGCTGGGGGC 597
Qy 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 598 AGCGGGGATGTTGGCGCCGCCAAGTATGACAGTGGTGGATCGGGCTTACCAAGTACGAC 657
Qy 141 SerGlyGluIleProArgGlyTrrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 658 AGTGGTGAGATCCCTCGGGGATACATCCCGTCACTCACTAACGCCAGATTTCCGGAGAA 717
Qy 161 IleProGlyAlaSerProAspHisHisMetSerProThrGlyAsnIleSerArgArg 180
Db 718 ATCCCTGGTCTCCCTCGACCATCATATGATGTCTCTACTGGGAACTTGGCAGGCGC 777
Qy 181 AlaProPheProTrrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
Db 778 GCCCATTTCCCTATATGAATCATTCATCAATCCGTCGAGGGAATTCCTCTGGTAGCGTT 837
Qy 201 GlyAsnValAlaTrrLysGluArgValAspGlyTrrLysMetLysGlnAspLysGlyAla 220
Db 838 GGGAAATGTTGCTGGAAGAGAGGGTTGATGGCTTGGAAATTCGAAGCAGGACAAAGGAA 897
Qy 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
Db 898 ATTCCCATGAGCATGGCACCAAGCATTTGCTCCCTCTGAGGGCCGGGTGTTGGTGATATT 957
Qy 241 AspAlaSerThrGluTrrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
Db 958 GATGCATCAACTGATTAACAATGGAAGATGCTCTATTAAACGATGAACTCGCCAGCCT 1017
Qy 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTrrArgMetValIleVal 280
Db 1018 CTATCTAGGAAAGTTCACCTTCTCCAGGATAAATCATACAGGATGGTCAATTGTG 1077
Qy 281 LeuArgLeuValValLeuSerIlePheLeuHisTrrArgLeuThrAsnProValArgAsn 300
```

Db	1078	CTACGATTGATTGTTCTTAAGCATCTTCTTGCACTACCGGATCACAAATCCTGTGCGTAAT	1137
QY	301	AlaTyrProLeuTyrPheLeuSerValIleCysGluIleTyrPheAlaLeuSerTyrIle	320
Db	1138	GCATACCCACCTGTGCTTCTATCTGTATATGTAGATCTGTTGCTCTTTCTCGGATA	1197
QY	321	LeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLeuAspArgLeuAla	340
Db	1198	TTGGATCAGTTTCCAAAGTGGTTTCCATCAACCCGGAGACTTACTTGTATAGATCGCA	1257
QY	341	LeuArgTyrAspArgGluGluProSerGlnLeuAlaValAspIlePheValSer	360
Db	1258	TTAAGCTATGACCGGAAAGTGAGCCATCTCAGTTGGCTGCTGTGAACATTTTGTCAGT	1317
QY	361	ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu	380
Db	1318	ACTGTGCAACCAATGAAGAGCCTCCTCTTGTTCATCTGCAATACCGTGTATCCATTCTC	1377
QY	381	AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet	400
Db	1378	GCTGTGGACTATCTCTGGATAAGGTCCTCTGTATGTATCTGTATGATGAGTGCATATG	1437
QY	401	LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTyrValProPheVal	420
Db	1438	CTGACATTTGATGCCTAGCTAGACTTCAGAGTTTGTCTAGAAAATGGGTGCCATTTGTT	1497
QY	421	LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr	440
Db	1498	AAGAAGTACCAACATTGAACCTAGAGCTCTGAATGGTACTTCTCCAGAAAATGATTAC	1557
QY	441	LeuLysAspLysValGlnProSerPheValLysAspArgArgAlaMetLysArgGluTyr	460
Db	1558	TTGAAGGACAAAGTGCACCTTCATTGTTAAAGACCGCGGCGCATGAAGAGAGATAT	1617
QY	461	GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu	480
Db	1618	GAAGAAATTCAAAATTAAGGTAAATGGCCTTGTGTGAAGGCACAAAAAGTCCTGAGGAA	1677
QY	481	GlyTyrIleMetGlnAspGlyThrProTyrProGlyValAsnThrArgAspHisProGly	500
Db	1678	GGATGGATCATGCAGATGGCACCATGGCCAGGAAACAATACCAGGGACCATCTTGGA	1737
QY	501	MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro	520
Db	1738	ATGATTCAGTTTCTCTGGTCCACAGTGGTGTCTGTACTAGGGTAATGAGTACCC	1797
QY	521	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly	540
Db	1798	CGTTTGGTCTATGTCTTCTGTAAGAACGTCCTGGATTCAGCATCACAGAAAGCTGGT	1857
QY	541	AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn	560
Db	1858	GCATGAATGCTCTTGTCCGCTCTCAGCTGTGCTTACCAGTGAACATACATGTTGAAT	1917
QY	561	LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu	580
Db	1918	CTTGATTTGATCACTACATCAACAACAGTAAGGCTCTCAGGGAAGCTATGTCTTCCT	1977
QY	581	MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly	600
Db	1978	ATGGATTCCTAACTAGGAGGAGTGCTGCTGTATGTTCAGTTTCCCCAGAGGTTGCGATGT	2037
QY	601	IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg	620
Db	2038	ATTGATAGGAATGATCGATATGCCAACAGGAACACCGTGTCTTTCGATATTAACTTGAGA	2097
QY	621	GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr	640
Db	2098	GGTCTTGATGGCATCCAGGACAGTTTATGTGGGCACCTGGCTGTCTTTTCAACAGACA	2157
QY	641	AlaIleTyrGlyTyrGluProProLysAlaLysLysProGlyPheLeuAlaSerLeu	660
Db	2158	GCTCTATATGGTTATGAGCCCCCAATTAAAGCAAAAGAGGGTGGTTTCTTGTTCATCACTA	2217
QY	661	CysGlyGlyLysLysAlaSerLysSerLysLysArgSerSerAspLysLysSer	680
Db	2218	TGTGTTGACGAGAGAGGAGCAATCAAGAAG--GGCTCAGACAAGAAAAGTCA	2274
QY	681	AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal	700
Db	2275	CAGAAGCATGTGCACAGTCTCTGTGCCAGTATTCAATCTTTGAAGATATAGAGGAGGTT	2334
QY	701	GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys	720
Db	2335	GAAGCGCTGGATTTGATGATGAGAAATCACTTTATGTCTCAATGAGCTTGGAGAAG	2394
QY	721	ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro	740
Db	2395	AGATTTTGCCAAATCTGAGCTTTTGTTCGATCCACTCTGATGGAATATGGTGGTCTCT	2454
QY	741	GlnSerSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGlyTyr	760
Db	2455	CAGTCTGGACTCCAGAATCTCTCTGAAGAAGAGCTATCCATGTATCAAGTTGTGGCTAC	2514
QY	761	GluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp	780
Db	2515	GAGGACAGATTGAATGGGAACCTGAGATTGGTGGATCTATGGTCTCTGACGGAGAT	2574
QY	781	IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys	800
Db	2575	ATTTCTACTGGTTCAAGATGCACGACGAGGCTGGCGTGCATCTACTGTCATGCCAAG	2634
QY	801	ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu	820
Db	2635	CGCGCGGCTTCAAGGATCGGCTCCCATCAATCTCTCAGACCGTCTGAACCCAGGTGCTC	2694
QY	821	ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGly	840
Db	2695	CGGTGGGTCTCGGTTCAGTGGAATTCCTTTTCAGCGGCATTTGCCCTATGGTACGGG	2754
QY	841	TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro	860
Db	2755	TACGAGGACCGCTGAGTTCTTGAGAGATTCCGCTACATCAACACCATCTACCCG	2814
QY	861	LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly	880
Db	2815	CTCAGTCCCTCCCGCTCCTCAATTTACTGTATCTCTGCTGCCATCTGCCGTGTACGGG	2874
QY	881	LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu	900
Db	2875	AAAGTTTCATCATCCAGAGATCAGCAACTTCGCTAGTATCTGGTTTCATCTCTCTTCATC	2934
QY	901	SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr	920
Db	2935	TCGATCTTCGCCACGGGTATCTCGAGATGAGTGGAGCGCGTGGGCATCGACGAGTGG	2994
QY	921	TyrArgAsnGluGlnPheTyrValIleGlyIleSerAlaHisLeuPheAlaValPhe	940
Db	2995	TGGAGGAAACGAGCAGTTCTGGGTTCATCGGAGCATCTCCGCCACCTCTTCGCCGTCTC	3054
QY	941	GlnGlyLeuLeuLysValIleAlaGlyIleAspThrAsnPheThrValThrSerLysAla	960
Db	3055	CAGGGCTCTCCTCAAGGTGTTGCCGTCATCGAACCACTTCCCGTTCACCTTCCAGGCC	3114
QY	961	AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeuIle	980
Db	3115	TCGGATGAAGACGGCGACTTCGCGGAGCTGTACATGTTTCAAGTGGACGACACTTCTGATC	3174
QY	981	ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla	1000
Db	3175	CGCGCCACCATCTGATCATCAACCTTGGTGGCGTGTGTGGCGGACATCTCTACGCC	3234
QY	1001	IleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPheAlaPheTyr	1020
Db	3235	ATCAACAGCGGGTACAGTCTGTGGGTCTGCTCTTCGGCAAGCTCTTCTTCGCTTCTCG	3294


```
QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
DB 1618 GAAGAAATTCAAAATTAGGTAATAGCCCTGTGTCTAAGGCACAAAAGTCCCTGAGGAA 1677
QY 481 GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 500
DB 1678 GGATGGATCATGCAAGATGGCACACATGGCCAGGAAAACAATACAGGAGCACCATCTGGA 1737
QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
DB 1738 ATGATTCAGGTTTCTTGGTCACAGTGTGTCTTGATCTAGAGGTAAATGAGCTACCC 1797
QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisIleLysLysAlaGly 540
DB 1798 CGTTGGTCTATGTTCTCGTGAAAACGTCCTGGATCCAGCATCACAAAGAAGCTGCT 1857
QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
DB 1858 GCCATGAATGCTCTTGTCCGCTCTCAGCTGTGCTTACCAATGGACAATACATGTTGAAT 1917
QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
DB 1918 CTTGATTGTGATTCATCATCAACACAGTAAGGCTCTCAGGGAAGCTATGTGCTTCCT 1977
QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
DB 1978 ATGGATCCTAACCCTAGGAGGAGTGTCTGTATGTTCAGTTTCCCAGAGGTTGATGCT 2037
QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPheAspIleAsnLeuArg 620
DB 2038 ATTGATAGGAATGATCGATATGCCAACAGAAACACCGTGTCTTTCGATATTAACCTTGA 2097
QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
DB 2098 GGTCTTGATGGCATCAAGACACAGTTATGTGGGCACCTGGCTGTGTTTCAACAGAAACA 2157
QY 641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
DB 2158 GCTCTATATGTTATGAGCCCCCAATTAAGCAAAAGAGGGTGTCTTGTTCATCACA 2217
QY 661 CysGlyGlyLysLysLysAlaSerLysLysLysArgSerSerAspLysLysLysSer 680
DB 2218 TGTGTGGCAGCAAGGAAGCAAAATCAAGAAG--GGCTCAGACAAGAAAAGTCA 2274
QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
DB 2275 CAGAAAGCATGTGCAGATTCTGTGCCAGTATTCAATCTTGAAGATATAGAGGAGGATT 2334
QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
DB 2335 GAAGGGCTGGATTGATGATGAGAAATCACTTCTTATGCTCAATGAGCTTGGAGAAG 2394
QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
DB 2395 AGATTGTGCCAAATCTGCAGCTTTTGTGTGCTCCACTCTGATGGAATATGCTGTCTCT 2454
QY 741 GlnSerSerThrProGlnSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
DB 2455 CAGTCTGGCAGCTCCAGAAATCTCTCTGAAAAGAGCTATCCATGTATAGTTGTGGCTAC 2514
QY 761 GluAspLysSerGluTrpGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
DB 2515 GAGGACAAGATTGAATGGGGAATCAGATTGGGTGGATCTATGTTCTGTGACGGAAGAT 2574
QY 781 IleLeuThrGlyPheLysMethIleAlaArgGlyTyrArgSerValTyrCysMetProLys 800
DB 2575 ATTCTCACTGGGTTCAAGATGCACGACGAGGCTGGCGTCCGATCTACTGTCATGCTAAG 2634
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
DB 2635 CGGCCGGCTTCAAGGGATCGGCTCCCATCTCTCAGACCGCTCTGAAACGAGGTGCTC 2694
```

```
QY 821 ArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyrGly 840
DB 2695 CGGTGGGCTCTCGGTTCAGTGGAAATCTTTTTCAGCGGCAATGGCCCTATGTTACGGG 2754
QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
DB 2755 TACGAGAGCGCTCAAGTCTCTTGAGAGATTTCGGCTACATCAACACCAACCATATACCG 2814
QY 861 LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
DB 2815 CTCAGTCTCCCTCCGCTCTCTATTACTGTATCTCTGCTGCCATCTGCTGCTACGGGG 2874
QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
DB 2875 AAGTTTCATCATCCAGAGATCAGCAACTTCGCTAGTATCTGTTTCTATCTCTCTTCATC 2934
QY 901 SerIlePheAlaThrGlyLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr 920
DB 2935 TCGATCTTTCGCCACCGGTATCTCGAGATGAGGTGAGCGGTGGGCATCGACGAGTGG 2994
QY 921 TrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaValPhe 940
DB 2995 TGGAGAACGAGCAGTCTTGGGTCTATCGAGGCATCTCGGCCCACTCTTCGCGCTCTTC 3054
QY 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
DB 3055 CAGGGCTCTCTCAAGGTGCTTCCGGCATCGACACCAACTTCACGCTCACCTCCAAGGCC 3114
QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
DB 3115 TCGGATGAAGACGGCGACTTCGGGAGCTGTACATGTTCAAGTGAGCAGACATCTCTGATC 3174
QY 981 ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
DB 3175 CGGCCCAACCATCTCTGATCATCAACCTGTTCGGGTGTGTCCGGCATCTCTACGCC 3234
QY 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
DB 3235 ATCAACAGCGGTACAGTCTGGGGTCCGCTCTTCGCAAGCTCTCTTCGCTCTCTGG 3294
QY 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
DB 3295 GTGATCGTTACCTGATCCGCTTCCTCAAGGCTCTCATGGGTGGGAGAACCGACCCCG 3354
QY 1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
DB 3355 ACCATCGTGTGTCTGGCGCATCTGCTGGCGTGTGATCTCTCTCTGCTGGGTTCGC 3414
QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
DB 3415 ATCGATCGGTTCAACAACCGCTCACTGGCCCGGATCTCGAAACGTTGGCATCAACTGC 3474
```

RESULT 5

US-10-209-059-21

; Sequence 21, Application US/10209059

; Publication NO. US2003016383A1

; GENERAL INFORMATION:

; APPLICANT: Dhugga, Kanwarpal S.

; APPLICANT: Wang, Haiyin

; TITLE OF INVENTION: Maize Cellulose Synthases and Uses

; TITLE OF INVENTION: Thereof

; FILE REFERENCE: 0864R2

; CURRENT APPLICATION NUMBER: US/10/209,059

; CURRENT FILING DATE: 2002-07-31

; PRIOR APPLICATION NUMBER: 60/096,822

; PRIOR FILING DATE: 1998-08-17

; PRIOR APPLICATION NUMBER: 09/371,383

; PRIOR FILING DATE: 1999-08-06

; PRIOR APPLICATION NUMBER: 09/550,483

; PRIOR FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 52

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 21

Db	1018	CTATCTAGGAAAGTTCCACCTTCCTCTCCAGGATAAATCCATACAGGATGGTCAATTGCT	1077
Qy	281	LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn	300
Db	1078	CTACGATTGATTGTTCTTAAGCATCTTCTTGCACTACCGGATCACAATCTCTGTGGTAAAT	1137
Qy	301	AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle	320
Db	1138	GCATACCACCTGTGGCTTCATCTGTTATATGTGAGATCTGGTTGCTCTTCTCTCGATAT	1197
Qy	321	LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla	340
Db	1198	TTGATCAGTTTCCAAAGTGGTTTCCAAATCAACCGCAGACCTTACCTTGATAGACTCGCA	1257
Qy	341	LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer	360
Db	1258	TTAAGGTATGACCGGAAGGTGAGCCATCTCAGTTGGCTGCTGTTGACATTTTGTCACT	1317
Qy	361	ThrValAspProLeuLysGluProProIleValIleThrAlaAsnThrValLeuSerIleLeu	380
Db	1318	ACTCTCACCAATGAGGAGCCTCTCTTCTTCTCACTGCCAATACCGTGCCTATCCATTCCTC	1377
Qy	381	AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet	400
Db	1378	GCVTGGACATCTCTGFGAATAGGTCTCTTGCTATGATCTGATGATGAGCTGCTATG	1437
Qy	401	LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal	420
Db	1438	CTGCATTTGATGCATGACTGCTGAGACTTCAGAGTTTGCTAGAAAATGGGTGCCATTTGTT	1497
Qy	421	LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr	440
Db	1498	AAGAAGTACAACTATGAACCTAGAGCTCTCTGAATGGTACTCTCCACAGAAAATGATTAC	1557
Qy	441	LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr	460
Db	1558	TTGAAGGACAAAGTGCACCCCTTCATTTGTTAAAGACCGCGGCCATGAAGAGAGAAATAT	1617
Qy	461	GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu	480
Db	1618	GAAGAAATCAAATTAGGGTAATGGCTTTGTTGCTTAAGGCACAAAAGTCCCTGAGGAA	1677
Qy	481	GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly	500
Db	1678	GGATGGATCATGCAAGATGCGACACCATGCGCAGGAAACAATACCAGGGACCATCCTGGA	1737
Qy	501	MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro	520
Db	1738	ATGATTTCAGGTTTTCTTGTGTCACAGTGGTGGTCTTGATACTCAGGGGTAAATAGCTACCC	1797
Qy	521	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly	540
Db	1798	CGTTTGGTCTATGTTCTCGTGAAAACGTCCTGGATCCAGCATCAACAAGAAGCTGGT	1857
Qy	541	AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn	560
Db	1858	GCCATGAATGCTCTTGTCCGCGTCTCAGCTGTGCTTACCAATGCAATACATGTTGAAT	1917
Qy	561	LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu	580
Db	1918	CTTGATTGTGATCACTACATCAACACAGTAAAGGCTCTCAGGGAAGCTATGTGCTTCCCTT	1977
Qy	581	MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly	600
Db	1978	ATGGATCCTTAACCTAGGAAGAGTGTCTGCTATGTTCAGTTTCCCCAGAGGTTTCGATGGT	2037
Qy	601	IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg	620
Db	2038	ATTGATAGGAATGATCGATATGCAACAGAAACACCGTGTGTTTTCGATATTAACTTGAGA	2097
Qy	621	GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr	640
Db	2098	GGTCTTCATGGCATCCAAGACACAGTTATGTGGCAGCTGGCTGTGTGTTTCAACAGAAC	2157

```
QY 641 AlalleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
Db 2158 GCTCATATGGTTATGAGCCCAATTAAGCAAAAGAGGGTGGTTCTTGTCATCACTA 2217
QY 661 CysGlyGlyLysLysLysAlaSerLysLysLysArgSerSerAspLysLysLysSer 680
Db 2218 TGTGGTGGCAGGAGCAAGGAGCAATCAAGAAG--GGCTCAGACAAGAAAAGTCA 2274
QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db 2275 CAGAGCATGTGGACAGTCTGTGGCAGTATTCAATCTTGAAGATATAGAGGAGGAGTT 2334
QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db 2335 GAAGCGCTGGATTGATGAGAAATCACTTCTATGTCTCAATGAGCTTGGAGAG 2394
QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
Db 2395 AGATTGGCCAATCTGCAGCTTTTGTGCTCCACTCTGATGGAATATGGTGTCTCT 2454
QY 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db 2455 CAGCTCTGCAGCTCCAGATCTCTCTGAAGAAGAGCTATCCATGTATAGTTGTGGCTAC 2514
QY 761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
Db 2515 GAGGACAGATTGAATGGGAACCTGAGATTGGTGGATCTATGGTCTGTGACGGAAGAT 2574
QY 781 IleLeuThrGlyPheLysMethHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
Db 2575 ATTCTCACTGGGTTCAAGATGACGACGAGGCTGGCGTGCATCTACTGATGCTTAAG 2634
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db 2635 CGCGCGGCTCTCAAGGGATCGGCTCCCATCAATCTCTCAGACCGTCTGAACCAAGTGTCT 2694
QY 821 ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGly 840
Db 2695 CGGTGGGCTCTCGTTCAGTGAATCTTTTCAGCGGGCATTTGCCCTATGTTACGGG 2754
QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db 2755 TACGAGGACGCTCAAGTCTCTGGAGAGATTGGCTATCATCAACACCACTATCCCG 2814
QY 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2815 CTCAGTCTCCCTCCGCTCTCTCAATTACTGTATCTCGCTGCGCATCTGCTGCTCAGGG 2874
QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
Db 2875 AAGTTCATCATCCAGAGATCAGCAACTCGCTAGTATCTGGTTCTCTCTCTTCATC 2934
QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr 920
Db 2935 TCGATCTTCGACACGGTATCTCGAGATGAGTGGAGCGGCGTGGGATCGACAGTGG 2994
QY 921 TrpArgAsnGluGlnPheTyrValIleGlyIleSerAlaHisLeuPheAlaValPhe 940
Db 2995 TGGAGAACGACGAGTCTTGGGTCTATCGAGGCACTTCGCCCACTCTTCGCCGTCTTC 3054
QY 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db 3055 CAGGCGCTCTCAAGGTCTTGGCGCATCGACCACTTCCACCTCACCTCCCAAGGCC 3114
QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeuLe 980
Db 3115 TCGATGAAGACGGGACCTTCGCGAGGTGTACATGTTCAAGTGCAGACACTTCTGTATC 3174
QY 981 ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db 3175 CGGCCCAACCACTCATCATCAACCTGTGCGGTTGTTGCCGCTATCTCTTCACGCC 3234
```

```
QY 1001 IleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPhePheAlaPheTyr 1020
Db 3235 ATCAACAGCGGTACCACTCGTGGGTCGGCTCTTCGCAAGCTCTTCTTCGCTTCTGG 3294
QY 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3295 GTGATCGTTCACCTGATCCGTTCTCAAGGCTCTCATGGGTCTCGCAGAAACCGCACCG 3354
QY 1041 ThrIleValIleValTyrAlaValLeuLeuAlaSerIlePheSerLeuLeuTyrValArg 1060
Db 3355 ACCATCGGTGTGTCTGGCGCATCTGTCGCTGATCTTCTCTGCTGTGGTTCGC 3414
QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3415 ATCGATCGTTCACCAACCGGCTCACTGGCCCGGATATCTCGAACGTGTGGCATCACTGC 3474

RESULT 6
US-10-160-719-29
; Sequence 29, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3551)
US-10-160-719-29

Alignment Scores:
Pred. No.: 0 Length: 3746
Score: 5421.50 Matches: 998
Percent Similarity: 96.67% Conservative: 46
Best Local Similarity: 92.41% Mismatches: 33
Query Match: 93.83% Indels: 3
DB: 15 Gaps: 2

US-09-900-237A-30 (1-1080) x US-10-160-719-29 (1-3746)
QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 321 ATGGAGGCGCAGCGGACGGGTGAAGTCGGGAGCGCGGTGGCGGACAGGTGTGCCAG 380
QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 381 ATCTGCGCGCAGCGGCTGGGCCACCGCGGAGGGGAGCGTCTTTCGCCGCTGCGACGTC 440
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 441 TCGGGTTCGCTGTCGCCCTTCCTACAGTACGAGCGGCGGACGAGCGGACGCGGCGG 500
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 501 TGCCCCCAGTGCAGACCAAGTACAAGCGCCACAAGGGGAGCGCCGCGCATCGTGGGAG 560
QY 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
```

Db 561 GAAGGAGCAGACTGATGCCGAT-----AGCGACTTCAATTACCTTGTCATCTGGCAAT 614
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
Db 615 GAGGACGAGAGAGAGATGGCCGACAGAAATGCCAGCTGGCGCATGAACGTTGGGGGC 674
QY 121 SerGlyAsnValGlyHisProLysTyAspSerGlyLulileGlyLeuSerLysTyAsp 140
Db 675 AGCGGGGATGTTGGTCGCCCCAGATATGACAGTGGCGAGATCGGGCTTTCACAGATATGAC 734
QY 141 SerGlyGluLeuProArgGlyTyValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 735 AGTGGCGAGATTCTCGGGGATATCCCATCAGTCACTAACACCGCAGATCTCAGAGAA 794
QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 795 ATCCCTGGTGTCTCCCTGACCATCATATGATGTCCCAATCGGGAACATTTGGCAAGCGT 854
QY 181 AlaProPheProTyValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
Db 855 GCTCCATTTCCCTATGTGAACCATTCGCCAAAATCCGTCAGGGAGTTCTCTGTAGCAAT 914
QY 201 GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyAla 220
Db 915 GCGAATGTTCCCTGGAAAGAGAGGTTGATGGCTGGAAAATGAAGCAGGACAAAGGGGACG 974
QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
Db 975 ATTTCCCATGACGAATGGCAACAAGCATTTGCTCCCTCTGAGGGTTCGGGTGTGTGATATT 1034
QY 241 AspAlaSerThrGluTyValAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
Db 1035 GATGCATCACTGATACACATGGAAGATGCCCTATTGACGCGAATCTCGACGCCT 1094
QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyArgMetValIleVal 280
Db 1095 CTATCTAGGAAAGTTCCACTTCTCTCCAGGATAAATCCATACAGATGGTCAATTGTG 1154
QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyArgLeuThrAsnProValArgAsn 300
Db 1155 CTGCGATGATGTTCTTAAGCATCTTCTGCACATACCGTATCAAAATCTCTGCGCAAT 1214
QY 301 AlaTyProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle 320
Db 1215 GCATACCATATAGGCTTCTATCTGTATATGTGAGATCTGGTTGCTCTTCTGTGGATA 1274
QY 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyLeuAspArgLeuAla 340
Db 1275 TTGGATCAGTTCCCTAAGTGGTTTCCAAATCAACCGGGAGAGCTACCTTGATAGGCTGGCA 1334
QY 341 LeuArgTyAspArgGluGlyGluProSerGlnLeuAlaValAspIlePheValSer 360
Db 1335 TTAAGSTATGACCGGAAGGTGAGCCATCTCAGTTGGCTGCTGTGATGATTTGCTCAGT 1394
QY 361 ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu 380
Db 1395 ACAGTCACCAATGAAGAGAGCTCTCTGTCTACTGCCAATACCGTCTATCATCTCT 1454
QY 381 AlaValAspTyProValAspLysValSerCysTyValSerAspAspGlyAlaSerMet 400
Db 1455 GCTGTGGATTACCTGTGGATAAGTCTCTTGCTATGTATCTGATGATGAGCTGGCATG 1514
QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
Db 1515 CTGACATTTGATGCACTAGCTGAGACTTCAGATTTGCTAGAAATGGGTACCAATTTGTT 1574
QY 421 LysLysTyAspIleGluProArgAlaProGluPheTyPhePheCysGlnLysIleAspTy 440
Db 1575 AAGAAGTACACATTTGAACCTTAGAGCTCTCGAATGGTACTTCTCCCAAGAAATTTGATTAC 1634
QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTy 460
Db 1635 TTGAAGGACAAAGTGCCACCTTCATTTGTTAAAGACCGCGGGCCATGAAGAGAAATAT 1694

QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
Db 1695 GAAGAATTCAAAGTATAGGTAAATGGCTTGTGTCTAAAGGCACAGAAAGTTCTTGAGAA 1754
QY 481 GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 500
Db 1755 GGATGGATCATGCAAGNATGGCACACCATGGCCAGGAAACAAATACCAGGACCATCTCTGGA 1814
QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
Db 1815 ATGATTCAGGTTTTCTTGGTCAAGTGGTGGCTTGATCTAGGCAATAGCTACCC 1874
QY 521 ArgLeuValTyValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 540
Db 1875 CGTTGGTCTATGTTCTCGTGAAGAGCTCTCGGATTCAGCATCAACAGAAAGCTGGT 1934
QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyMetLeuAsn 560
Db 1935 GCCATGAATGCTCTTGTTCGTGCTCAGCTGTGCTTACCAATGGACAAATACATGTTGAAT 1994
QY 561 LeuAspCysAspHisTyIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
Db 1995 CTTGATTTGTCATCACTACATTAACAACAGTAAGGCTCTCAGGGAAGCTATGTGCTTCTCT 2054
QY 581 MetAspProAsnLeuGlyProGlnValCysTyValGlnPheProGlnArgPheAspGly 600
Db 2055 ATGGACCTTAACCTAGGAGAGGTGCTGCTCAGTCCAGTTTCCCGAGATTCGATGGC 2114
QY 601 IleAspArgAsnAspArgTyValAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
Db 2115 ATTGACAGGAATGATCGATATGCCACAGAAACACCGTGTTCGATATTAATTGATGAGA 2174
QY 621 GlyLeuAspGlyIleGlnGlyProValTyValGlyThrGlyCysValPheAsnArgThr 640
Db 2175 GGTCTTGATGGCATCCAAGGACAGATTATGTCCGAACCTGGCTGTGTTTCAACCGGAACA 2234
QY 641 AlaIleTyGlyTyGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
Db 2235 GCTCTATATGTTATGAGCCCCCAATTAAGCAAGAAGGGTGGTTTCTTGTCTATCACTA 2294
QY 661 CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer 680
Db 2295 TGTGGCGGTAGGAAGAGGCAAGCAATCAAGAG---GGCTCGACACAGAGAAGTCG 2351
QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db 2352 CAGAAGCATGTGGACAGTTCTGTGCCAGTATTCAACCTTGAAGATATAGAGGAGGAGTT 2411
QY 701 GluGlyAlaGlyPheAspAspGluLysSerValIleuMetSerGlnMetSerLeuGluLys 720
Db 2412 GAAGGGCGCTGGATTTGACGAGAGAAATCACTTCTATGTCTCAATAGCGCTGGAGAG 2471
QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyGlyGlyValPro 740
Db 2472 AGATTTGGCCAGTCCGACAGGTTTGTTCCTCCACTCTGATGGATGATGGTGTCTCT 2531
QY 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTy 760
Db 2532 CAGTCCGCAACTCCGGAGTCTCTTCTGAAAGAGCTATCCATGTATTAAGCTGTGGCTAT 2591
QY 761 GluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyGlySerValThrGluAsp 780
Db 2592 GAGGACAAGACTGAATGGGAACTGAGATCGGGTGGATCTACCGTTCTGTGACAGAGAC 2651
QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyCysMetProLys 800
Db 2652 ATTCTACCGGATTCAGATGCAAGCGGAGGCTGGCGGTCTGATCTACTGATGCCCAAG 2711
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db 2712 CGGCCAGCTTCAAGGGGTCTGCCCCCATCAATCTTTTCGACCGCTCTGAACAGGTGCTC 2771

```
Qy 821 ArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyrGly 840
Db 2772 CGGTGGGTCTTGGTTCGGTGGAGATCTCTTCACCGGCACTGCCCTCTGTGTACGGC 2831
Qy 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db 2832 TACGAGGCGCGCTCAAGTTCCTGGAGAGATTCGGGTACATCAACACCACTCTACCGC 2891
Qy 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2892 CTCAGTCCATCCCTCTCTCATCTACTACTGATCTGCGCGGCATCTGCTGTCTACCGGA 2951
Qy 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeu 900
Db 2952 AAGTTCATCATCTCCAGAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTCTCATC 3011
Qy 901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrp 920
Db 3012 TCGATCTTTCGCCACCGGCATCTCGAGATAGAGTGGAGCGGGTGGGCATCGACGAGTGG 3071
Qy 921 TrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaValPhe 940
Db 3072 TGGAGAACAGCAGCATCTTGGGTGATCGGGGCACTCTCGGCCACTCTTTCGCCGTGTTC 3131
Qy 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db 3132 CAGGGCTTGTCTCAAGGTCTGGCGGCATCGACACCACTTCACCGTCACTCCCAAGGCC 3191
Qy 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrLeuLeuIle 980
Db 3192 TCGGACGAGGAGCGGCACTTCGCGAGGTGTACATGTTCAAGTGGACGCGTCTCTGATC 3251
Qy 981 ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db 3252 CGCCCCACCATCTGATCATCACTGATGTCGGCGTCTGTCGGCGCATCTCTCTACGCC 3311
Qy 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
Db 3312 ATCAACAGCGGATACAGTCTGGGGCGCGCTCTTCGGCAAGCTCTTCTTGGCTTCTGG 3371
Qy 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3372 GTCACTGTCCACTGATCCGCTTCTCAAGGCGCTCATGGGCGAGCAACCGCACCCCG 3431
Qy 1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
Db 3432 ACCATCTGCTGCTCTGGGCCATCTGCTGGCTCCATCTCTCTGCTGTGGGTTCGC 3491
Qy 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3492 ATCGACCCCTTCAACACCGCGTCACTGGCGCGGATACCCAGACGTGTGGCATCAACTGC 3551
```

RESULT 7

```
US-10-160-719-49
; Sequence 49, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helencjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 49
```

```
; LENGTH: 3746
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (321)...(3449)
US-10-160-719-49
Alignment Scores:
Pred. No.: 0 Length: 3746
Score: 5421.50 Matches: 998
Percent Similarity: 96.67% Conservatives: 46
Best Local Similarity: 92.41% Mismatches: 33
Query Match: 93.83% Indels: 3
DB: 15 Gaps: 2
US-09-900-237A-30 (1-1080) x US-10-160-719-49 (1-3746)
Qy 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 321 ATGGAGGCGCAGCGGCAGCGGTGAAGTCGGGAGCGCGGTGGCGGACAGGTGTGCCAG 380
Qy 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysaspVal 40
Db 381 ATCTCGCGCGCAGCGGTGGGCACCAAGCGCGGAGGGGAGCGTCTTCGCCCTCGCACGTC 440
Qy 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 441 TCGCGGTTCCTGGTGTGCGGCCCTCTGCTACGATGACGAGCGGACGAGCGGACGCGGCG 500
Qy 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 501 TGCCCCCAGTGCAGACCAAGTACAAAGCGCCACAGGGGAGCGCGCGATCGCTGGGAG 560
Qy 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 561 GAAGGAGACGACACTGATGCGCAT-----AGCGACTTCAATTACCTTGCATCTCGCAAT 614
Qy 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
Db 615 GAGGACCAAGAGCAGAGAGATTGCCACAGATTGCCAGCTGCGCATGAACTGGCGGCGC 674
Qy 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 675 AGCGGGGATGTGGTTCGCCCCCAAGTATCACAGTGGCGAGATCGGGCTTACCAAGTATGAC 734
Qy 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 735 AGTGGCGAGATTCTCTCGGGGATACATCCCATCAGTCACTAAACAGCCAGATCTCAGAGAA 794
Qy 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 795 ATCCCTGGTGTCTTCCCTGACCATCATATGATGATGTCCTCCCACTGGGAAACATTGGCAAGCGT 854
Qy 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
Db 855 GCTCCATTTCCCTATGTGAACCACTTCGCAAAATCCGTCAAGGGAGTCTCTGTGTAGCATT 914
Qy 201 GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGluAspLysGlyAla 220
Db 915 GGGAAATGTGCTCGAAAGAGAGGGGTTCATGCTGGAAAAATGAACAGCAGGCAAGGGGACG 974
Qy 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
Db 975 ATTCCTCATGACGAATGGCACAAGCATCTCTCTCTGAGGGTTCGGGTGTGGTATATT 1034
Qy 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
Db 1035 GATGCATCAACTGATTACAAACATGAGAGATGCTTTATTGAACGACGAAATCTGACAGGCT 1094
Qy 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
Db 1095 CTATCTAGGAAGTTCCCACTCTCTCTCCAGGATAAATCCATATCAGGATGGTTCATTGTG 1154
```



```
Db 3312 ATCAACAGCGGATACAGTCGTGGGGCGCGCTCTTCGGCAAGCTCTTCTTGCCCTTCGG 3371
Qy 1021 VallileValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3372 GTCAFCGTCCACCTGTACCGCTTCCTCAAGGCGCTCATGGGCGAGCAAGACCGACCCCG 3431
Qy 1041 ThrileValIleValTrrAlaValLeuLeuAlaSerIlePheSerLeuLeuTrrValArg 1060
Db 3432 ACCATCGTCGTCTGGGCCATCTCTGGGGTCCATCTTCTTCTGTGGGTTCGC 3491
Qy 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3492 ATCGACCCCTTCAACACCGCGTCACTGGCCCGGATACCCAGACGTGGCATCAACTGC 3551

RESULT 8
US-10-627-132-9
; Sequence 9, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; PRIOR FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-9

Alignment Scores:
Pred. No.: 0 Length: 3773
Score: 5421.50 Matches: 998
Percent Similarity: 96.67% Conservative: 46
Best Local Similarity: 92.41% Mismatches: 33
Query Match: 93.83% Indels: 3
DB: 13 Gaps: 2

US-09-900-237A-30 (1-1080) x US-10-627-132-9 (1-3773)
Qy 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 338 ATGAGGCGGACGCGGACGGCGTGAAGTCGGGGAGCGCGGTGGCGGACAGGTGTGCCAG 397
Qy 21 IleCysAlaAspGlyLeuGlyThrThrLeuaspGlyAspValPheThrAlaCysaspVal 40
Db 398 ATCTGCGCGGACGGCGTGGGCACCGCGGAGGGGACGCTCTTCGCCGCGCTGCGACGTC 457
Qy 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 458 TCGGGTTCCTCGGTGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 517
Qy 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 518 TGCCCCCAGTGCACAGCAAGTACAAAGCGCCACAAAGGGGAGCGCCGCGATCGGTGGGAG 577
Qy 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 578 GAAGGAGACGACACTGATGCCGAT-----AGCGACTTCAATTACTTGCATCTGCGCAAT 631
```

```
Qy 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrrArgMetAsnThrGlyGly 120
Db 632 GAGGACACAGAGCAGAGCAAGATTGCCACAGAAATCGCAGCTGGCGCATGAACGTTGGGGC 691
Qy 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 692 AGCGGGATGTTGTCGCCCCCAAGTATGACAGTGGCAGATCGGCTTACCAGATATGAC 751
Qy 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 752 AGTGGCAGAGATTCCTCGGGGATACATCCCATCAGTCACTAACAGCCAGCATCAGAGAA 811
Qy 161 IleProGlyAlaSerProAspHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 812 ATCCCTGGTGTCTCCCTGACCATCATATGATGTCCTCCCACTGGGAACATTTGGCAAGCGT 871
Qy 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
Db 872 GCTCCATTTCCCTATGTGAACCATTCGGCAAAATCCGTCAAGGGAGTTCTCTGTGTAGCAT 931
Qy 201 GlyAsnValAlaTrrLysGluArgValAspGlyTrrLysMetLysGlnAspLysGlyAla 220
Db 932 GGGAAATGTCCTGGAAAGAGAGGGTTGATGCTCGAAAAATGAACAGACAGCAAGGGGAG 991
Qy 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
Db 992 ATTCCTCATGACGAATGGCACAAGCATTCCTCCCTGAGGGTGGGGTGTGTGTATATT 1051
Qy 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
Db 1052 GATGCATCAACTGATTACAAATGGAAGATGCTTTATTGAACAGACGAACATCGACAGCT 1111
Qy 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
Db 1112 CTATCTAGAAAGTTCCACTCTTCCTCCAGGATAAATCCATACAGATGGTCAATTGTG 1171
Qy 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
Db 1172 CTGCGATTGATTGTTCTAAGCATCTTCTTGCACTACCGTATCACAAATCCTGTGCGCAAT 1231
Qy 301 AlaTyrProLeuTrrLysLeuSerValIleCysGluIleTrrPheAlaLeuSerTrpIle 320
Db 1232 GCATACCCATTTATGGCTTCTATCTGTTATATGTGAGATCTGCTTTGCTTTCGTGGATA 1291
Qy 321 LeuAspGlnPheProLysTrrPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
Db 1292 TTGGATCAGTTCCCTPAAGTGGTTTCCAAATCAACCGGGAGACGCTACTTGTAGAGCTGCA 1351
Qy 341 LeuArgTyrAspArgGluGlyProSerGlnLeuAlaAlaValAspIlePheValSer 360
Db 1352 TTAAGGTATGACCGGGAAGGTGAGCCATCTCAGTTGGCTGCTGTGTGACATTTTCGTCA 1411
Qy 361 ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu 380
Db 1412 ACAGTCGACCCCAATGAAGGAGCCTCCTTGTGCTACCTGCAATACCGTCTATCCATTCTT 1471
Qy 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
Db 1472 GCTGTGGATTACCCCTGTGGATAAGGTCTCTTGTCTAATGATCTGTAGATGGAGCTCGATG 1531
Qy 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrrValProPheVal 420
Db 1532 CTGACATTTGATGCATAGCTGAGACTTCAGACTTTGTAGAAAAATGGTACCATTTGTT 1591
Qy 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
Db 1592 AAGAGTACACATTTGAACCTTAGAGTCTCCTGAATGGTACTTCTCCCAAAAAATGATTAC 1651
Qy 441 LeuLysAspLysValGlnProSerPheValLysAspArgArgAlaMetLysArgGluTyr 460
Db 1652 TTGAAGGACAAAGTGCACCTTCATTGTTAAAGACCGCGGGCCATGAAGAGAGAAAT 1711
Qy 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
```

```
Db      1712 GAAGAAATCAAAGTTAGGCTTAAGTCTGTGCTAAGGCACAGAAAGTTCTCTGAGAA 1771
QY      481 GlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly 500
Db      1772 GGATGGATCATGCAAGATGACACACCATGCGCAGGAAACAATACCAAGGACCATCTGGA 1831
QY      501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
Db      1832 ATGATTCAGGTTTCTTGTGTCTCAGTGGTGGCTTGTATCTAGGCAATAGCTATCCC 1891
QY      521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 540
Db      1892 CGTTTGTCTATGTTCTCTGTAAGAGCGTCTGGATTCAGCATCACAGAAGCTGGT 1951
QY      541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
Db      1952 GCCATGAATCTCTTGTGTCTCTCAGCTGTGCTTACCAATGACAAATACATGTTGAAT 2011
QY      561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
Db      2012 CTTGATTTGATCACTTACATTAACAAGTAAGCTCTCAGGAAGCTATGTGCTTCCTT 2071
QY      581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
Db      2072 ATGGACCTTAACCTAGGAAGAGTGTCTGCTACGTCCAGTTTCCCCAGAGATTTCGATGGC 2131
QY      601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
Db      2132 ATTGACAGGAATGATCGATATGCAACAGGAACACACCGTGTGTTTTCGATATTAACCTTGAGA 2191
QY      621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
Db      2192 GGTCTTGATGGCATCCAAAGCAGACAGTTTATGCGAACTGGCTGTGTTTTCACCCGAACA 2251
QY      641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
Db      2252 GCTCTATATGTTATGAGCCGCCCAATTAAGCAGAGAGGGTGGTTCTTGTCTACACTA 2311
QY      661 CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerAspLysLysLysSer 680
Db      2312 TGTGGCGGTAGGAAGGAGCAAGCAATCAAGAAG---GGCTCGGCAAGAAGAAGTCG 2368
QY      681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db      2369 CAGAAGCATGTGGACAGTCTCTGTCAGTATTAACCTTGAAGATATAGAGGAGGAGTT 2428
QY      701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db      2429 GAAGGGCTGGATTTGACGACGAGAAATCACTTCTTATGCTCAAAATGAGCCTGGAGAAG 2488
QY      721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
Db      2489 AGATTTGGCCAGTCCGACAGCTTGTGTGCTCCACTCTGATGGAGTATGTGTGTTCCT 2548
QY      741 GlnSerSerThrProGluSerLeuLeuLysLysAlaIleHisValIleSerCysGlyTyr 760
Db      2549 CAGTCCGCACATCGGAGTCTCTTCTGAAGAAGCTATCCATGATTAAGCTGTGGCTAT 2608
QY      761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
Db      2609 GAGGACAAGACTGNATGGGNACTGAGATCGGGTGGATCTACGGTCTGTGACAGAAGAC 2668
QY      781 IleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetProLys 800
Db      2669 ATTCACCGGAATTCAGAGTGCACGCGGAGGCTGGCGGTGCGATCTACTGATGCCCAAG 2728
QY      801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db      2729 CGGCCAGCTTTCAGGGGTCTGCCCTCCCAATCTTTCCGACCGCTCTGAACCCAGGTCTC 2788
QY      821 ArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTrpGly 840
```

```
Db      2789 CGGTGGGCTCTTGGGTCCGTGGAGATCTCTTTCAGCGCGGCACTGCCCTGTGGTACGGC 2848
QY      841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db      2849 TACGGAGGGCGGTCTCAAGTTCTCTGGAGAGATTCGGGTACATCAACACCATCTACCCG 2908
QY      861 LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db      2909 CTCAGTCCATCCGCTTCTCATCTACTGATCTCTGCCGCCATCTGTCTGTCTACCGGA 2968
QY      881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeu 900
Db      2969 AAGTTTCATCATCCAGAGATCAGCAACTTCGCCAGCATCTGGTTTCATCTCCCTCTTCATC 3028
QY      901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTyr 920
Db      3029 TCCATCTTCGCCACGGGCATCTCTGGAGATAGGTGGAGCGGGTGGGCATCGCAGAGTGG 3088
QY      921 TrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaValPhe 940
Db      3089 TGGAGGAACGAGCAGTTCTGGGTGATCGGGGGCATCTCCGCGCACCTCTTCGCGGTTC 3148
QY      941 GlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db      3149 CAGGGCTGTCTCAAGGTGCTGGCGGCATCGACACCAACTTCACCGTCACCTCCAGGGCC 3208
QY      961 AsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
Db      3209 TCGAGCAGAGACGGCGACTTCGCGGAGCTGTATCATGTTCAAGTGGAGCGCGCTCTGATC 3268
QY      981 ProProThrThrIleLeuIleLeuAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db      3269 CCGCCACACCATCTGATCATCACTGTCGGGTGTCGCGCGCATCTCTCTAGGCC 3328
QY      1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
Db      3329 ATCAACAGCGGATACAGTGTGGGGCCGCTCTTGGCAAGCTCTTCTTCTCGCTCTCGG 3388
QY      1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db      3389 GTCATGCTCACCTGTACCCGTTCTCAAGGGCTCATGGCGGCGCATGGCGAGAACCGCCCG 3448
QY      1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
Db      3449 ACCATGTCGTCTCTGGGCCATCTCTGTCGGCTGCATCTTCTCTCTGTGTGGTTCGC 3508
QY      1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db      3509 ATCGACCCCTTCACCCCGCGTCACTGGCGCGGATACCCAGACGTTGTGGCATCAACTGC 3568

RESULT 9
US-10-209-059-9
; Sequence 9, Application US/10209059
; Publication No. US2003016383A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Haiyin
; APPLICANT: Dhugga, Kanwarpal S.
; TITLE OF INVENTION: Maize Cellulose Syntheses and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R2
; CURRENT APPLICATION NUMBER: US/10/209,059
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3773
; TYPE: DNA
```



```
Db 2312 TGTGGCGGTAGGAGAGAGGACGAATCAAGAG--GCCTCGGACAGAGAGACTCG 2368
Qy |||||:::|||||
Db 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Qy |||||:::|||||
Db 2369 CAGAAGCATGTGCACAGTTCTGTGCCAGTATTCAACCTTGAAGATATAGAGGAGGAGTT 2428
Qy |||||:::|||||
Db 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Qy |||||:::|||||
Db 2429 GAAGGGCGCTGGATTGACGACGAGAAATCACTTCTTATGTCTCAAAATGAGCCCTGGAGAG 2488
Qy |||||:::|||||
Db 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
Qy |||||:::|||||
Db 2489 AGATTGGCCAGTCCGCGAGGTTTGTGCTCTCACTCTGATGGAGTATGGTGTCTCT 2548
Qy |||||:::|||||
Db 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Qy |||||:::|||||
Db 2549 CAGTCGCAACTCCGGAGTCTCTTCTGAAAGAGCTATCCATGTTATAAGCTGTGGCTAT 2608
Qy |||||:::|||||
Db 761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrPheIleTyrGlySerValThrGluAsp 780
Qy |||||:::|||||
Db 2609 GAGGACAAGACTGAATGGGGAACAGATCGGGTGGATCTACGGTCTGTGACAGAAAGC 2668
Qy |||||:::|||||
Db 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
Qy |||||:::|||||
Db 2669 ATTCTACCGGATTCAAGATGCAGCGCGAGGCTGGCGGTGATCTACTACTGCATGCCCAAG 2728
Qy |||||:::|||||
Db 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Qy |||||:::|||||
Db 2729 CGGCCAGCTTCAAGGGGTCTGCCCCCATCAATCTTTCGAGCCGTCTGAACCAAGTCTC 2788
Qy |||||:::|||||
Db 821 ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGly 840
Qy |||||:::|||||
Db 2789 CGGTGGGCTCTTGGGTCCGGAGATCCTCTTCAGCCGGCACTGCCCCCTGTGTGTACGGC 2848
Qy |||||:::|||||
Db 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Qy |||||:::|||||
Db 2849 TAGCGAGGCGGCTCAAGTTCCTGGAGAGATTCGGGTACATCAACACCAACCATACCCG 2908
Qy |||||:::|||||
Db 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Qy |||||:::|||||
Db 2909 CTCAGTCCATCCGCTTCTCATCTACTGCTATCTGCGCGCACTCTCTGCTCACCGGA 2968
Qy |||||:::|||||
Db 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleThrPheIleAlaLeuPheLeu 900
Qy |||||:::|||||
Db 2969 AAGTTCAATTCAGAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTCTTCATC 3028
Qy |||||:::|||||
Db 901 SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr 920
Qy |||||:::|||||
Db 3029 TCGATCTTCGCCAGCGCATCTCGAGATGAGGTGGAGCGGGTGGGCATCGACGAGTGG 3088
Qy |||||:::|||||
Db 921 TrpArgAsnGluGlnPheThrValIleGlyCylIleSerAlaHisLeuPheAlaValPhe 940
Qy |||||:::|||||
Db 3089 TGGAGGAACAGAGAGTTCGGGTGATCGGGGATCTCCGCGCACTCTTCGCGGTTC 3148
Qy |||||:::|||||
Db 941 GlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Qy |||||:::|||||
Db 3149 CAGGGCTGTCTAAGGTGCTGGCGGATCGACACCACTTCACCGTCACCTCCAGGCC 3208
Qy |||||:::|||||
Db 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeuIle 980
Qy |||||:::|||||
Db 3209 TCGGACGAGGAGCGGACTTCGCGAGCTGTACATGTTCAAGTGGACGACGCTCTGATC 3268
Qy |||||:::|||||
Db 981 ProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Qy |||||:::|||||
Db 3269 CCGCCCAACCACTCTGATCATCAACTGTGCGCGTCTGTCGCGGCATCTCTACGCC 3328
Qy |||||:::|||||
Db 1001 IleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPheAlaPheThr 1020
Qy |||||:::|||||
Db 3329 ATCAACAGCGGATACCAAGTGTGGGGCCGCTCTTCGCGCAAGCTCTTCTTCGCTTCGG 3388
Qy |||||:::|||||
Db 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Qy |||||:::|||||
```

```
Db 3389 GTCATCGTCACCTGTACTACCGCTTCTCAAGGGGCTCATGGGACGAGAACCGCACCCCG 3448
Qy |||||:::|||||
Db 1041 ThrIleValIleValThrAlaValLeuLeuAlaSerIlePheSerLeuLeuTyrValArg 1060
Qy |||||:::|||||
Db 3449 ACCATCGTCGTCGTGGGCATCTCTGGCGTCATCTTCTCTTCTGTGGGTTCG 3508
Qy |||||:::|||||
Db 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Qy |||||:::|||||
Db 3509 ATCGACCCCTTACCACCCCGCTACTGGCCCGATACCCAGACGTTGGCATCACTGC 3568

RESULT 10
US-10-160-719-5
; Sequence 5, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 3773
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (338)....(3568)
US-10-160-719-5

Alignment Scores:
Pred. No.: 0 Length: 3773
Score: 5421.50 Matches: 998
Percent Similarity: 96.67% Conservative: 46
Best Local Similarity: 92.41% Mismatches: 33
Query Match: 93.83% Indels: 3
DB: 15 Gaps: 2

US-09-900-237A-30 (1-1080) x US-10-160-719-5 (1-3773)
Qy 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 338 ATGAGGGCGACGCGGACCGGCGTGAAGTCGGGGAGCGCGGTGGCGACAGGTGTGCCAG 397
Qy 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 398 ATCTGCGCGACGCGGTGGGCACCAACCGCGAGGGGACGCTCTTCGCGCTGCGAGCTC 457
Qy 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db 458 TGGCGGTTTCCGGTGTGCCGCCCTGCTACGAGTACGAGCGCAAGGACGCGCAGCGCG 517
Qy 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 518 TGCCCCCAGTGCAGACCAAGTCAAGCGCCCAAGGGGAGCCCGCGATCCCGTGGGGAG 577
Qy 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 578 GAAGGACGACACTGATGCGCAT-----AGCGACTTCAATTACCTTGATCTGGCAAT 631
Qy 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
Db 632 GAGGACCAAGCAGAAAGATTGCCGACAGAAATGCCGAGCTGGCGCATGAACGTTGGGGGC 691
```

QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
DB : : : : :
DB AGCGGGGATGTGGTCGCCCCCAAGATGACAGTGGGAGATCGGGCTTACCAAGTATGAC 751
QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
DB : : : : :
DB AGTGGCGAGATTCTCTCGGGGATACATCCCATCAGTCACTAAACAGCAGATCTCAGAGAA 811
QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArg 180
DB : : : : :
DB ATCCCTGGTGCTTCCCTCGACCATCATATGATGTCCTCCCACTGGGAAACATTGGCAAGCGT 871
QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
DB : : : : :
DB GCTCCATTCCCTATGTGAACCATTCGCCCAATCCGTCAAGGGAGTCTCTCGTAGCATTT 931
QY 201 GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyAla 220
DB : : : : :
DB GGGAAATGTGCTGGAAAGAGAGGGTTGATGGCTGGAAATGAACGACAGGAGGAGG 991
QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
DB : : : : :
DB ATTCCCATGACGAATGGCACAGCATTTGCTCCCTCTGAGGGTCGGGGTGTGGTGATATT 1051
QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
DB : : : : :
DB GATGCATCAACTGATTACAAATGGAAGATGCTTTATTGAACGACGAAATCGACAGCCT 1111
QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
DB : : : : :
DB CTATCTAGGAAAGTTCCACTCTCTCCCTCCAGGATAAATCCATACAGGATGGTCAATTGT 1171
QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
DB : : : : :
DB CTGCGAATTGATTGTTCTAAGCAATCTCTTGCACTACCGTATCAAAATCTGTGGCAAT 1231
QY 301 AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle 320
DB : : : : :
DB GCATACCCATTATGGCTTCTATCTGTATATGTGAGATCTGGTTTGTCTTTCGTGGATA 1291
QY 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgIleAla 340
DB : : : : :
DB TTGGATCATCTCCCTAAGTGGTTTCCAAATCAACCGGAGACGTACTTGATAGGTGGCA 1351
QY 341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
DB : : : : :
DB TTAAGGTATGACCGGGAAGGTGAGCCATCTCAAGTTGGCTGTCTGTGACATTTCTGCAGT 1411
QY 361 ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu 380
DB : : : : :
DB ACAGTCGACCAATGAAGGAGCCTCCTCTTGTCACTGGCCAAATACCGTGTATCCATTCTT 1471
QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspGlyAlaSerMet 400
DB : : : : :
DB GCTGTGGATTACCCCTGGGATTAAGGTCTCTGCTATGATATCATGATGGAGTCGGATG 1531
QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
DB : : : : :
DB CTGACATTGTATGCACTAGCTGAGACTTCAGAGTTTGTAGAAATGGGTACCATTTGTT 1591
QY 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
DB : : : : :
DB AAGAAGTACAACTTGAAGCTAGAGCTCTGTAATGGTACTTCTCCGAGAAATTCATTATC 1651
QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAtqAlaMetLysArgGluTyr 460
DB : : : : :
DB TTGAAGGACAAAGTGCACCCCTTCAATTTGTTTAAAGACCGCGGGGCCATGAAGAGAAAT 1711
QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
DB : : : : :
DB GAAGAATTCAAAGTTAGGGTAAATGGCCTTGTGTCTAAGGCACAGAAAGTTCTCTGAGAA 1771
QY 481 GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 500

DB 1772 GGATGGATCATGCAAGATGGCACCATGGCCAGGAAACAAATACAGGGACCATCTCTGGA 1831
QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
DB : : : : :
DB ATGATTTCAGGTTTCCCTTGGTCACAGTGGTGGCCTTGTACTAGAGGGCAATGAGTACCC 1891
QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 540
DB : : : : :
DB CTTTGGTCTATGTTTCTCGTGAAGAAGCGTCTCGATTTCCAGCATCACAGAAAGCTGGT 1951
QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
DB : : : : :
DB GCCATGAATGCTCTTGTTCGTCTCAGCTGTGCTTACCAATGGACAATACATGTTGAAT 2011
QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
DB : : : : :
DB CTTGATTGTGATCACTACATTAAACAACAGTAGGCTCTCAGGGAAGCTATGTGCTTCCT 2071
QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
DB : : : : :
DB ATGGACCTTAACCTAGGAAGGAGTGTCTGTACGTCCAGTTTCCCCAGAGATTTCGATGCG 2131
QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
DB : : : : :
DB ATTGACAGGAATGATCGATATGCCAACAGGAACACCGTGTCTTCGATATTAACTTGAGA 2191
QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
DB : : : : :
DB GGTCTGTGATGGATCCAAAGGACCATTTATGTCCGAACTGGCTGTGTTTCAACCGAAC 2251
QY 641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
DB : : : : :
DB GCTCTATATGTTATGAGCCCCCAATTAAAGCAGAAAGAGGGTGGTTTCTTGTTCATCATA 2311
QY 661 CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer 680
DB : : : : :
DB TGTGCGGTAGGAAGAGCAAGCAAAATCAAAGAAG--GGCTCGCAAGAAGAAGTCG 2368
QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGluVal 700
DB : : : : :
DB CAGAAGCATGTGACAGATTCTGTGCCAGTATTCAACCTTGAAGATATAGAGGGAGATT 2428
QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
DB : : : : :
DB GAAGCGCTGGATTGACGACGAGAAATCACTTCTTATGTCTCAATGAGCCTGGAGAAG 2488
QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
DB : : : : :
DB AGATTGGCCAGTCCGACGCGTTTGTGCTCCACTCTGATGGAGTATGGTGTGTTCT 2548
QY 741 GlnSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
DB : : : : :
DB CAGTCCGCACTCCGAGTCTCTTCTGAAGAAGCTATCCATGTTATAGCTGTGGCTAT 2608
QY 761 GluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGluAsp 780
DB : : : : :
DB GAGGACAAAGACTGAATGGGAACTCAGATCGGTGGTGAICTACGGTCTGTGTGACAGAGAC 2668
QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetProLys 800
DB : : : : :
DB ATTCTACCGGATTCAAGATGCACGCGAGGCTGCGGTGCATCTACTGATGCCCAAG 2728
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
DB : : : : :
DB CGGCAGCTTTCAAGGGTCTGCCCCCATCAATCTTTCCGACCGCTGNACCAGGTGCTC 2788
QY 821 ArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyrGly 840
DB : : : : :
DB CGGTGGGTCTTGGTCCGTGGAGATCCTCTTCAGCGGCACTGCCCTCTGTGTGACGCG 2848
QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860

Db 2849 TACGAGGGCGGCTCAAGTTCTCGGAGATTCCGGTACATCAACACCATCTACCCG 2908
Qy 861 LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2909 CTCACGTCCTACCGCTCTCATCTACTGATCTCGCCGCCATCTGTCTGTCTCACCGGA 2968
Qy 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleThrPheIleAlaLeuPheLeu 900
Db 2969 AAGTTCACTATCCAGAGATCAGCACTTCGCCAGCATCTGGTTCATCTCCCTCTTCATC 3028
Qy 901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrp 920
Db 3029 TCGATCTTCGCCAGGGCATCTCGAGATGAGGTGGAGCGGGTGGGCGATCGACGAGTGG 3088
Qy 921 TrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaIleThrPheAlaValPhe 940
Db 3089 TGGAGGAACAGCAGATTCTGGGTGATCGGGGCGATCTCCGCGCATCTCCGCGTGTTC 3148
Qy 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db 3149 CAGGGCTGTCTCAAGGTGCTGGCGGCGATCGACACCACTTCACCGTCACCTCCAGGCC 3208
Qy 961 AsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrLeuLeuIle 980
Db 3209 TCGACGAGGACGGGACTTCGCGGACCTGTACATGTTCAAGTGGAGCGCTCTCTGATC 3268
Qy 981 ProProThrThrLeuIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db 3269 CCGCCACCACTCTGATCATCAACTGTGTGGCGTGTGGCGGCGATCTCTCTAGGCC 3328
Qy 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPheAlaPheTrp 1020
Db 3329 ATCAACAGCGGATACCATGTGTGGGGCCGCTCTTCGGCAAGCTCTCTTCGCGCTTCGG 3388
Qy 1021 ValIleValIleLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3389 GTCATCGTCACCTGTACCCGCTTCTCAAGGCGCTCATGGGCGAGACCGCACCCCG 3448
Qy 1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
Db 3449 ACCATCGCTCGTCTGGGCGATCTCTGCGCGTCCATCTCTCTCTGCTGTGGGTTCG 3508
Qy 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3509 ATGACCCCTTCACACCGCGTCACTGCGCCGATACCCAGACGTGTGGCATCAACTGC 3568

RESULT 11

US-10-627-132-13
; Sequence 13, Application US/10627132
; Publication No. US20040068767A1
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal S.
; APPLICANT: Wang, Haiyin
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864R3
; CURRENT APPLICATION NUMBER: US/10/627,132
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/096,822
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 09/371,383
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 09/550,483
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 10/209,059
; PRIOR FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 3704
; TYPE: DNA
; ORGANISM: Zea mays
US-10-627-132-13

Alignment Scores:
Pred. No.: 0 Length: 3704
Score: 5149.00 Matches: 948
Percent Similarity: 94.08% Conservative: 69
Best Local Similarity: 87.70% Mismatches: 58
Query Match: 89.11% Indels: 6
DB: 13 Gaps: 6

US-09-900-237A-30 (1-1080) x US-10-627-132-13 (1-3704)

Qy 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 272 ATGACGGC---GGCGAGCGCCACGAATTCGGGGAAGCATGTGCGCGGAGGTGTGCAG 328
Qy 21 IleCysAlaAspGlyLeuGlyThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 329 ATCTGCGGCGACGGCGTGGCGACCGCGCGCGACCGCCCTTTCACCGCTGCGAGCTC 388
Qy 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGlyThrGlnAla 60
Db 389 TGGGCTTCCCGTGTGCGCCCATGCTACGAGTACGAGCGCAAGGACCGCACCGCGC 448
Qy 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 449 TGCCCGCAGTGCAGACTAAGTACAGCGCCACAAAGGGAGCCACCAGTACACGGTGAG 508
Qy 81 GluGlyAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
Db 509 GMAATGAGGATGTGGATGTGACGATGTGAGTACACTACCACTACCAAGCATCTGGCAAC 568
Qy 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
Db 569 CAGGATCAGAAAGCAAAAGATTGCTGAGAGAAATGCTCACTTGGCGGGAACAACCTACGTCG 628
Qy 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 629 AGT---GATATTGGCTGGCTAAGTATGACAGCGGTCAAAATTTGGCGCATGGGAGTATGAC 685
Qy 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 686 AGTGGTCAAGATCCCTCGTGTGATATATCCCTCACTAACTCATAGCCAGATCTCAGAGAG 745
Qy 161 IleProGlyAlaSerProAsnHisHisMetSerProThrGlyAsnIleSerArgArg 180
Db 746 ATTCTCGAGCTTCCCTCGAT---CATATGATGTCCTCTGTTGGAAACAATTGGCAGCGGT 802
Qy 181 Ala---ProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySer 199
Db 803 GGCATCAATTTCTTATGTAATCATTTCTCAAAACCATCGAGGGAGTTCTCCGGTAGC 862
Qy 200 IleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGly 219
Db 863 CTGGCAATGTTGCATGAAAGAGAGGAGGTGGATGGATGGAATAAAG---GATAAAGGT 919
Qy 220 AlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAsp 239
Db 920 GCAATTCCTATGACCAATGGAAAGCAAGCATTTGCTCCATCAGAAGGGCGTGGAGTTGCTGAT 979
Qy 240 IleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGln 259
Db 980 ATTGATGCTTCTACTGATTATACATGGAAGATGCTTACTGATGATGATGAACCTCGGCAA 1039
Qy 260 ProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIle 279
Db 1040 CCTCTACTAGAAAGTGCCAAATTCCTTCATCCAGATAAATCCGTACAGAAATGGTTCATT 1099
Qy 280 ValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArg 299
Db 1100 GTGCTACGTTTGGCTGTCTTCTATGTCATATTTCTGCGCTACCGGTATCACACATCTGTGAAC 1159
Qy 300 AsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrp 319

Db 1160 AATGCATATCCACTGTGGCTTTATCCGTCATATGTGAGATCTGGTTGCTTTGTCTCTGG 1219
Qy 320 IleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeu 339
Db 1220 ATTTTGGATCAGTTCCCAAGTGGTCCCAATCAACCGTGAACATACCTTGTATAGACTG 1279
Qy 340 AlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheVal 359
Db 1280 GCTTTAAGGTATGACCGAAGGTGAACCATCTCAATTAGCTCCTGTGATATTTTGTGTC 1339
Qy 360 SerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIle 379
Db 1340 AGTACTGTGGATCCCAATGAAGAGGCTCCTCTGTCTACTGCAAAATACGTGTCTTCCATC 1399
Qy 380 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer 399
Db 1400 CTTGCTGTGCAATATCCGGTTGACAGGATCTTGTCTATGTTTCGATGATGAGCTGCT 1459
Qy 400 MetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPhe 419
Db 1460 ATGCTGACTTTTGATGCTCTCTGAAACTTCAGAGTTTGTAGAAAATGGGTTCGGTTC 1519
Qy 420 ValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAsp 439
Db 1520 TGTAAAGAGTACAAATAGAGCTTAGGCCCGCGAATGGTACTTTGCTCAGAAAATTGAT 1579
Qy 440 TyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGlu 459
Db 1580 TACTTGAAGACAAAGTTCAACCTCATTTGTGAAGAAGACGGCGGCATGAAGAGAGAA 1639
Qy 460 TyrGluGluPheLysIleAArgIleAsnAlaLeuValSerLysAlaLeuLysValProGlu 479
Db 1640 TATGAAGAAATTCAAAGTTTCGTATCAATGCTCTGTAGCCAGGCACAAAAAGTTCCCGAG 1699
Qy 480 GluGlyTyrIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro 499
Db 1700 GAGGATGGATCATCAGATGGTACACCTGGCCCTGGGAAACAATACTAGGGACCATCT 1759
Qy 500 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeu 519
Db 1760 GGAATGATCAGTTTTCCTGGTACAGTGGAGGGCTTGCGTTGAAGGCATGAACTT 1819
Qy 520 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAla 539
Db 1820 CCTCGTTGGTTTATGTGCTCGTCAAAAACTCTCGGATTCACACATCAACAAGAGCT 1879
Qy 540 GlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeu 559
Db 1880 GGTGCCATGAATGCATTTGCTGTATCAGCTGCTTACTTAATGGGCAATACATGTTG 1939
Qy 560 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPhe 579
Db 1940 AATCTTGATGTGACCACTACATCAATATAGCAAGGCTCTTCGAGAAGCTATGTGCTTC 1999
Qy 580 LeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAsp 599
Db 2000 CTTATGGACCCAAACCTAGGAAGGAATGCTGTGTATGTCCAATTTCTCAGAGGTTTGAT 2059
Qy 600 GlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPheAspIleAsnLeu 619
Db 2060 GGTATGTATGAAGATGACCATATGCAAAACAGGAACACTGTGTTTTCGATATTAACCTG 2119
Qy 620 ArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArg 639
Db 2120 AGAGCTCTTGACGCATTCAGGGCCAGTTTATGTGGAACTGGTTGTGTGTTTAACAGA 2179
Qy 640 ThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSer 659
Db 2180 ACGGCTTATATGTTATGAGCCTCCAGTCAAGAAAAAAGCCAGGCTTCTTCTCTTCG 2239
Qy 660 LeuCysGlyGlyLysLysAlaSerLysSerLysLysArgSerAspLysLysLys 679
Db 2240 CTTTGTGGGGAGGAGAAAAAGACGTCAAAATCTAAGAAAG---AGCTCGAAAAAGAGAG 2296

Qy 680 SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGly 699
Db 2297 TCACATAGACACCGACAGACTTCTGTACCAAGTATTAACTCGAAGATATAGAGGAGG 2356
Qy 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db 2357 ATTGAAGGTTCTCAGTTTGATGATGAGAAATCGCTGATTATGCTCTCAATGAGCTGGAG 2416
Qy 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyVal 739
Db 2417 AAGAGATTGGCCAGCTCCAGTGTGTTTGTAGCTCTACTCTGATGGAATATGTTGTGT 2476
Qy 740 ProGlnSerThrProGlnSerLeuLysGluAlaIleHisValIleSerCysGly 759
Db 2477 CCACAACTGCAACTCCAGAGTCTCTCTGAAGAAGACTTATCATGTCACTGCTGGC 2536
Qy 760 TyrGluAspLysSerGluTrpGlyThrGluIleGlyTyrIleTyrGlySerValThrGlu 779
Db 2537 TATGAGGACAAAACCTGACTGGGAACTGAGATTGGGTGGATCTATGTTCTCTTACAGAA 2596
Qy 780 AspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetPro 799
Db 2597 GACATTCTCACCGGATTCAAGATGCATGCTCGAGCTCGCGATCAATCTACTGCATGCT 2656
Qy 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
Db 2657 AAGCGACACAGCTTTCAGGGATCTGCTCTATCAACCTTTCGATCGTTTGAATCAAGTG 2716
Qy 820 LeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyr 839
Db 2717 CTTCCGTGGGCTCTTGGTTCCATTGAAATTCCTTTTCAGCAGGCACTGTCCTATATGTTAT 2776
Qy 840 GlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
Db 2777 GCCTATGAGGCGCGCTTAAATTCCTGGAGAGATTGCTTATATCAACACAAATTTAT 2836
Qy 860 ProLeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
Db 2837 CCACCTCATCAATCCCGCTCTCTGTACTGTCATATTCGACGAGTTGTCTTCTCACT 2896
Qy 880 GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPhe 899
Db 2897 GGGAAATTCATCATCCCAAGATTAGTAACCTAGAGAGTGTGGTTTATATCGCTCTTT 2956
Qy 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGlu 919
Db 2957 ATCTCAATCTTTGCCACTGGTATCCTTGAGATGAGGTGGGTGGTTCATGATGAA 3016
Qy 920 TrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaVal 939
Db 3017 TGGTGGAGAACAGCAGCTTCTGGGTCAATGGTGGTATTTCTGCGCATTTATTTGCCGTC 3076
Qy 940 PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 959
Db 3077 TTCAGGGTCTCTGAAGGTGCTTGTGATCGATCGACACAGCTTCACTGTCACTCTAAG 3136
Qy 960 AlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeu 979
Db 3137 GCCACTGACGAAGAAGGTGATTTTGGCGAGCTCTACATGTTCAAGTGGACAAACGCTTCG 3196
Qy 980 IleProProThrThrIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyr 999
Db 3197 ATCCCAACCAACTATTTTTCATCATCAACTGCTGGCTGGCTGGCTGGCATTTCTTAC 3256
Qy 1000 AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3257 GCAATCAATACGGGTTTACAGTCATCGGGAACCTCTTTTCGGGAAAGCTCTTCTTTCGCTC 3316
Qy 1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGluAsnArgThr 1039
Db 3317 TGGGTGATTGTCCACTGTATCCCTTCTCAAGGCGCTCTAAGGGAGACGACAAACCGCACG 3376

Db 1700 GAGGATGGATCATGCAAGATGGTACACCTTGGCTGGGAAACAATACTAGGACCATCTCT 1759
Qy 500 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeu 519
Db 1760 GGAATGATTCAGGTTTCCTGGGTACACGTGGAGGCTTGAGCTTGAAGCAATGAACATT 1819
Qy 520 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisIshisLysAla 539
Db 1820 CCTGGTTGGTTATGTCTCTGTAAGAAACGCTCGGATTCACATCAACAGAGGCT 1879
Qy 540 GlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeu 559
Db 1880 GGTGCATCAATGCACTTGTCTGTATCAGCTGTCTTACTAATGGGCAATACATGTTG 1939
Qy 560 AsnLeuAspCysAspHisfYrIleAsnSerLysAlaValArgGluAlaMetCysPhe 579
Db 1940 AATCTTGATGTGTGACCACTACATCAATAATAGCAGGCTCTTCGAGAAGCTATGTGCTTC 1999
Qy 580 LeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAsp 599
Db 2000 CTTATGGACCCAAACCTAGGAAGGATGCTGTATGTCCAAATTCCTCAGAGGTTTGAT 2059
Qy 600 GlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeu 619
Db 2060 GGTATTGATAGGAATGACCGATATGCAACAGCAACACTGTGTTTTTCGATATTAACCTG 2119
Qy 620 ArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArg 639
Db 2120 AGAGGTCTTGACGGCAATTCAAGGCCAGTTTATGTGGAACTGGTGTGTGTTTAACAGA 2179
Qy 640 ThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSer 659
Db 2180 ACGGCTTATATGTTATGAGCTCCAGTCAAGAAAAAAGCCAGGCTTCTTCTTCG 2239
Qy 660 LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLys 679
Db 2240 CTTTGTGGGGGAGAAAAAGACGTCAAAATCTAAGAAG--AGCTCGGAAAAGAGAAG 2296
Qy 680 SerAsnLysHisValAspSerValProValPheAsnLeuGluAspIleGluGly 699
Db 2297 TCACATAGACGACGACAGTCTCTGACCAGATTTTAACTCGAAGATATAGAGGAGG 2356
Qy 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db 2357 ATTGAAGGTTCTCAGTTTCATGATGAGAAATCGCTGATTTATGTCATAATGAGCTTGGAG 2416
Qy 720 LysArgPheGlyGlnSerAlaPheValAlaSerThrLeuMetGluTyrGlyGlyVal 739
Db 2417 AAGAGATTTGGCCAGTCCAGTGTGTTTGTAGCCTCTACTCTGATGGAATATGGTGTGT 2476
Qy 740 ProGlnSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGly 759
Db 2477 CCACAACTCGAACTCCAGAGTCTCTTGAAGAAGCTATTTCATGTCATCAGCTGTGGC 2536
Qy 760 TyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGlu 779
Db 2537 TATGAGCAAAAACACTGACTGGGAACTGAGATTGGGTGATCTATGCTGTCTGTTACAGAA 2596
Qy 780 AspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetPro 799
Db 2597 GACATTCTCACCAGATTCAGATGCAATGCTCGAGGCTGGCGATCAATCTACTGCAATGCT 2656
Qy 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
Db 2657 AAGCGACCAAGCTTCAAGGAACTGCTCTCACTCAACCTTCGGATCGTTTGAATCAAGTG 2716
Qy 820 LeuArgTTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyr 839
Db 2717 CTTCTGGTGGCTCTTGGTTCCATTGAAATCTTTTCAGCAGGCAATGTCCTATATGAT 2776
Qy 840 GlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
Db 2777 GGTATGGAGGCGGCTTAAATTCCTGGAGAGATTTGCTTATATCAACAAATTTAT 2836

Qy 860 ProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
Db 2837 CCATCTACATCAATCCGCTCTCTGTACTGATATTCGATATTCGACAGCTTTGTCTTCTCAT 2896
Qy 880 GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPhe 899
Db 2897 GGGAAAGTTTCATCATCCAAAGATTAGTAACCTAGAGAGTGTGGTTTATATCGCTCTTT 2956
Qy 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGlu 919
Db 2957 ATCTCAATCTTTGCCACTGGTATCTTGATAGAGTGAGTGAGTGTTCGCAATTGAAGAA 3016
Qy 920 TrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaVal 939
Db 3017 TGGTGGAGAACGACGAGTTCCTGGGTCAATGGTGTATTTCTGCGCAATTTATTCGCCGTC 3076
Qy 940 PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 959
Db 3077 TTCCAGGGTCTCTGAAAGTGTCTGTGATTCGACACGAGCTTCACCTGCACCTCAAG 3136
Qy 960 AlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeu 979
Db 3137 GCCACTGACGAAGAAGGTGATTTGCCGAGCTCTACATGTTCAAGTGGACACCGCTTCG 3196
Qy 980 IleProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyr 999
Db 3197 ATCCCAACCAACCACTATTTTGATCATCAACCTGGTGGGTGGTGGTGGTGGTGGTGGT 3256
Qy 1000 AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3257 GCAATCAATAGCGGTATCCAGTCAATGGGAGCCTCTTTTCGGGAAAGCTCTTCTTGGCTTC 3316
Qy 1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
Db 3317 TGGGTGATGTCCACTGATACCCCTTCCTCAAGGGCTCATGGGAAAGCAGAACCCGACG 3376
Qy 1040 ProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1059
Db 3377 CCGACCATGTGTTGTCCTGGGCTATCTCTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3436
Qy 1060 ArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
Db 3437 CGTATCGATCCATTCAACCCCGGCTCACTGGGCTGATATCGGAAATGTGGCATCAAC 3496
Qy 1080 Cys 1080
Db 3497 TGC 3499
RESULT 13
US-10-160-719-57
; Sequence 57, Application US/10160719
; Publication No. US20030167528A1
; GENERAL INFORMATION:
; APPLICANT: Kanwarpal S. Dhugga
; APPLICANT: Timothy G. Helentjaris
; APPLICANT: Benjamin A. Bowen
; APPLICANT: Xun Wang
; TITLE OF INVENTION: Maize Cellulose Synthases and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0864
; CURRENT APPLICATION NUMBER: US/10/160,719
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/371,383A
; PRIOR FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: US 60/096,822
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 57
; LENGTH: 3704
; TYPE: DNA
; ORGANISM: Zea mays

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (272) ....(3499)
US-10-160-719-57

Alignment Scores:
Pred. No.: 0 Length: 3704
Score: 5149.00 Matches: 948
Percent Similarity: 94.08% Conservative: 69
Best Local Similarity: 87.70% Mismatches: 58
Query Match: 89.11% Indels: 6
DB: 15 Gaps: 6

US-09-900-237A-30 (1-1080) x US-10-160-719-57 (1-3704)

QY 1 MetAspGlyAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
DB 272 ATGGACGGC---GGCGACGCCACGAATTCGGGAAGCATGTGGCGCGCAGGTGTGCCAG 328
QY 21 IleCysAlaAspGlyLeuGlyThrLeuAspGlyAspValPheThrAlaCysAspVal 40
DB 329 ATTCGGCGCACCGCGTGGGCACCGCGCGACGGCGACCGACCTCTTCACCGCGCTGCGACGTC 388
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGlyThrGlnAla 60
DB 389 TGGGGTTCCTCCGTGTGCCGCCCATGCTACGAGTACGAGCGCAAGGACGCCACCGCGCG 448
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
DB 449 TGGCGCGAGTGCAGACTAAGTCAAGCGCCACAAAGGGAGGCCACCGAGTACACGGTGAG 508
QY 81 GluGlyAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
DB 509 GAAATAGGAGTGTGGATGTGACGATGTGAGTACTACAACTACCAAGCATCTGGCAAC 568
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
DB 569 CAGGATCAGAGCAAGAAGATGTCTGAGAGATGTCTACTTGGCGGCAAACTACAGTGGC 628
QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
DB 629 AGT---GATATTGSCCTGGCTAAGTATGACAGCGGTGAATTTGGCGATGGGAAGTATGAC 685
QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
DB 686 AGTGGTGAGATCCCTCGTGTGATATATCCCGTCACTAACTCATAGCCAGATCTCAGGAGAG 745
QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
DB 746 ATTCTGGAGCTTCCCTCGAT---CATATGATGTCTCTGTGGGAACATTTGGCAGCGGT 802
QY 181 Ala---ProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySer 199
DB 803 GGACATCAATTTCTTATGTAATCAATCTCCAAACCCATCGAGGGGAGTCTCCGGTAGC 862
QY 200 IleGlyAsnValAlaIleProLysGluArgValAspGlyTyrLysMetLysGlnAspLysGly 219
DB 863 CTTGGCAATGTTGCATCGAAAGAGGGGTGGATGGATGGAAATGAAG---GATAAAGGT 919
QY 220 AlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAsp 239
DB 920 GCATTCCTATGACCAATGGAACAGCATGTCTCCATCAGAGGGCGGTGGAGTGTCTGAT 979
QY 240 IleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGln 259
DB 980 ATTGATGCTTCTACTGATTATAACATGGAGATGCCCTTACTGATGATGAACCTCGGCAA 1039
QY 260 ProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIle 279
DB 1040 CCTCTATCTAGAAAGTGCCAAATTCCTTCATCCAGAAATAAATCCGTACAGAAATGGTCAAT 1099
QY 280 ValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArg 299

```


Db 429 CAGGATCAGAAAGATTGCTGAGAGAAATGCTCACCTGGCGGACAAACTCACGTTGC 488
Qy 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluileGlyLeuSerLysTyrAsp 140
Db 489 AGT---GATATTGGCTGGCTTAAGTATGACAGCGGTGAAATTGGGCATGGGAAGTATGAC 545
Qy 141 SerGlyGluileProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 546 AGTGGTGAGATCCTCTGTGATATATCCCTGCTACTAACTCATGACCCAGATCTCAGGAGAG 605
Qy 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 606 ATTCCTGGAGCTTCCCTGTAT---CATATGATGTCTCTGTGGGAAACATTGGCGGCGGT 662
Qy 181 Ala---ProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySer 199
Db 663 GGACATCAATTTCTTATGTAAATCATTTCCAAACCCATCGAGGGAGTTCTCCGGTAGC 722
Qy 200 IleGlyAsnValAlaTrpLysGluAtrValAspGlyTrpLysMetLysGlnAspLysGly 219
Db 723 CTTGGCAATGTTGCATGGAAGAGAGGTGGATGGATGGAAATGAAG---GATAAAGGT 779
Qy 220 AlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAsp 239
Db 780 GCAATTCCTATGACCAATGGAACAAGCATTTGCTCCATCAGAAAGGGCGTGGAGTTGCTGAT 839
Qy 240 IleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGln 259
Db 840 ATTGATGCTTCTACTGATTATGACATGGAAGATGCCTTACTGAATGATGAACCTCGGCAA 899
Qy 260 ProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIle 279
Db 900 CCTCTATCTAGAAAAGGCCAATTCCTTCATCCAGAAATAATCCGTACAGAAATGGTCATT 959
Qy 280 ValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArg 299
Db 960 GTGCTAGCTTTGGCTGTTCTATGTCATATTTCTGGCTACCGCTATCACACATCCCTGTGAAC 1019
Qy 300 AsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluileTrpPheAlaLeuSerTrp 319
Db 1020 AATGCATATCCACTGTGGCTTTTATCGTCATATGTGAGATCTGGTTGGTTGCTCTCTGG 1079
Qy 320 IleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeu 339
Db 1080 ATTTTGATCAGTTCCTCCAAAGTGTGCTCCCAATCAACCGTGAACATACCTTGATAGACTG 1139
Qy 340 AlaLeuArgTyrAspArgGlyGluProSerGlnLeuAlaValAspIlePheVal 359
Db 1140 GCTTTAAGGTATGACCGAGAAGGTGAACCATCTCAATTAGCTCTGTTGATATTTTGTGTC 1199
Qy 360 SerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIle 379
Db 1200 AGTACTGTGATCCAAATGAAGGACCTCTCTGTCTCATCTGCAATACTGTCTTCCATC 1259
Qy 380 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer 399
Db 1260 CTTGCTGTGATATCCGGTTGACAAGGTATCTTGTGATGTTTCGGATGATGGAGCTGCT 1319
Qy 400 MetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPhe 419
Db 1320 ATGCTGACTTTGATGCTCTCTCGAACTTCAGAGTTTGTCTAGAAATGGGTTCCGTTTC 1379
Qy 420 ValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAsp 439
Db 1380 TGTAAGAAGTACAACATAGACCTTAGGGCCCGGAATGGTACTTTGCTCAGAAAAATTGAT 1439
Qy 440 TyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGlu 459
Db 1440 TACTTTGAAAGCAAAAGTTTCAAACTCATTTGTGAAAGAAACCCCGGCGCATGAAGAGAGAA 1499
Qy 460 TyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGlu 479
Db 1500 TATGAAGNAITCAAAGTTCGTATCAATGGTCTTTGTAGCCAGGCGACAAAAAGTTCCCGAG 1559

Qy 480 GluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro 499
Db 1560 GAGGGATGGATCATGCAAGATGTATACACCTTGGCTGGGAACAATACTAGGGACATCTCT 1619
Qy 500 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeu 519
Db 1620 GGAATGATTCAAGTTTCTCTGGGTCAAGTGGAGGGCTTCAGCGTTGAAGCAATGAATTT 1679
Qy 520 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAla 539
Db 1680 CCTCGTTTGGTTTATGTGTCTCGTGAACAAACGCTCTGGATTCCAACATCAAGAAAGCT 1739
Qy 540 GlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeu 559
Db 1740 GGTGCCATGAATGCATTTGTTGTGTATCAGCTGTCTTACTAATGGGCAATACATGTTG 1799
Qy 560 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPhe 579
Db 1800 AATCTTGATTGTGACCATACATCAATAATAGCAAGGCTCTTCGAGAAGCTATGTGCTTC 1859
Qy 580 LeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAsp 599
Db 1860 CTTATGACCCAAACCTTAGGAAGAAATGCTGTATTGTCCAATTTCTCAGAGGTTTGTAT 1919
Qy 600 GlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeu 619
Db 1920 GGTATTGATAGGAATGACCATATGCAACAGGAACACTGTGTTTTTCGATATTAACTTG 1979
Qy 620 ArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArg 639
Db 1980 AGAGTCTTGACGGCATTCAAGGGCCAGTTTATGTGGAACTGGTTGTGTGTTTAAACAGA 2039
Qy 640 ThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSer 659
Db 2040 ACGGCTTATATGTTATGAGCTCCAGTCAAGAAAAAAGCCAGGCTTCTCTCTCTCG 2099
Qy 660 LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLys 679
Db 2100 CTTTGTGGGGAAGGAAAAAGCTCAAAATCTAAGAAG---AGCTCGAAAAAGAGAG 2156
Qy 680 SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGluGly 699
Db 2157 TCATATAGACACGACAGACAGTTCTGTACCACTATTTAATCTCGAAGATATAGAGAAAGG 2216
Qy 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db 2217 ATTTGAAGGTTCTCAGTTTGTATGATGAGAAATCGCTGATTATGTCTCAATATGAGCTTGAG 2276
Qy 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyVal 739
Db 2277 AAGAGATTTGGCAGTCCAGTGTGTTTGTAGCCCTCTACTCTGATGGAATATGGTGGTGT 2336
Qy 740 ProGlnSerSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGly 759
Db 2337 CCACAACTGCAACTCCAGAGTCTCTCTGAAAGAACTATTATCATCATCAGCTGTGGC 2396
Qy 760 TyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIleTyrGlySerValThrGlu 779
Db 2397 TATGAGGACAAAACACTGCTGGGAACTGAGATTGGGTGGATATATGTTCTCTGTATACAGAA 2456
Qy 780 AspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArgSerValTyrCysMetPro 799
Db 2457 GACATTTCCCGGATTCAGAGTCAAGTCTCGAGGCTGGCGATCAATCTACTGTATGCCT 2516
Qy 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
Db 2517 AAGGACACAGCTTTCAGAGGATCTGTCTCTATCAACCTTTCCGATCCTTTGAATCAAGTG 2576
Qy 820 LeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTrpTyr 839
Db 2577 CTTTCGGTGGCTCTTGTTCCATTGAAATCTTTTTCAGCAGGCAATTTGTCCTCATATGTAT 2636

```
QY 840 GlyTyrGlyGlyArgLeuLeuPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
Db 2637 GCCTATGAGGCGCGCTAAATTCCTGGAGAGATTTCCTTATCAACACAACTTTAT 2696
QY 860 ProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
Db 2697 CCACCTCACATCAATCCGCTCCTCTGTAATGCTGATATTCGCCAGAGTTTGTCTTTCACT 2756
QY 880 GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPhe 899
Db 2757 GGAAGATTCATCATCCCAAGATTAGTACCTAGAGAGTGTGTGTTTATATCGCTCTTT 2816
QY 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGlu 919
Db 2817 ATCTCAATCTTTGCCACTCGTATCCTTTGAGATGAGTGGAGTGTGTCGCAATTGATGAA 2876
QY 920 TrpTrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaVal 939
Db 2877 TGGTGGAGAACGACGAGTTCCTGGGTCAATTTGGTGTATTTCTGCCGCAATTTATTTGCCGTC 2936
QY 940 PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 959
Db 2937 TTCCAGGGTCTCTGAGGTGTTTGGTGTATTCGATCGACAGCTTCACTGTCACTCTAAG 2996
QY 960 AlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeu 979
Db 2997 GCCACTGACGAAAGAGTGATTTTGGCGAGCTTACATGTTCAAGTGGACAAACGCTTCTG 3056
QY 980 IleProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlyThrSerTyr 999
Db 3057 ATCCACCAACACATATTTTGTATCATCAACCTGGTTCGGGTGCTGCTGGCAATTTCTTAC 3116
QY 1000 AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3117 GCATTCATAGCGGTACAGTCATCGGGACCTCTTTCCGGAGACTCTTCTTCGCTTC 3176
QY 1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
Db 3177 TGGGTGATTTGCACCTGTACCCCTTCTCAAGGGCTCATGGGAAAGCAGAACCGCAGC 3236
QY 1040 ProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1059
Db 3237 CGACCATTTGTCGTGTCTGGGCTATCTCTTCTGCTGATCTTTCTTCTGATGGGT 3296
QY 1060 ArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
Db 3297 CGTATCGATCCATTACACACCCGGGTCTACTGCCCCGTATATCGCAAAATGGGCATCAAC 3356
QY 1080 Cys 1080
Db 3357 TGC 3359
```

```
RESULT 15
US-10-437-963-39762/c
; Sequence 39762, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 39762
; LENGTH: 4264
```

```
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MPT4530_4326C.1
US-10-437-963-39762
```

```
Alignment Scores:
Pred. No.: 0 Length: 4264
Score: 5107.50 Matches: 944
Percent Similarity: 93.94% Conservatives: 64
Best Local Similarity: 87.98% Mismatches: 51
Query Match: 88.40% Indels: 15
DB: 17 Gaps: 4
```

US-09-900-237A-30 (1-1080) x US-10-437-963-39762 (1-4264)

```
QY 9 LysSerGlyArgHisGlyAlaGlyAspValCysGlnIleCysAlaAspGlyLeuGlyThr 28
Db 3431 AAGTCGGGAGGACGCGGAGCGCGCGTCCAGATCTTAGGTGACGGAG--GGC--- 3377
QY 29 ThrLeuAspGlyAspValPheThrAlaCysAspValCysArgPheProValCysArgPro 48
Db 3376 -----GACGTCTGGGGTTCCCGGTGCGCGGCCA 3347
QY 49 CysTyrGluHisGluArgLysGluGlyThrGlnAlaCysLeuGlnCysLysThrLysTyr 68
Db 3346 TGCTACGAGTACGAGCGCAAGGACGCGCGCGTCCCGCAGTGCACGACCAAGTAC 3287
QY 69 LysArgHisArgGlySerProAlaIleArgGlyGluGluGlyAspAspThrAspAlaAsp 88
Db 3286 AAGCCCAACAAAGGGAGCGCGCGTACTTGGGGATGAAAGCGATGATGTTTCATCGGAT 3227
QY 89 AspGlySerAspPheAsnTyrProAlaSerGlyThrGluAspGlnLysGlnLysIleAla 108
Db 3226 GATGCTAGTGTGATGTAATCTTCCAACTCCGCAACAGGACCAATAGCACAAGATGTC 3167
QY 109 AspArgMetArgSerTrpArgMetAsnThrGlyGlySerGlyAsnValGlyHisProLys 128
Db 3166 GAGAGGATGCTCACCTGGCGCATGAATCTGGGAGGAATGATGATATTGCTCATTAAG 3107
QY 129 TyrAspSerGlyGluIleGlyLeuSerLysTyrAspSerGlyGluIleProArgGlyTyr 148
Db 3106 TATGATAGTGTGAGATTGGTTCATCCCAAGTATGACAGTGGTGAATCCCTCGCATATAT 3047
QY 149 ValProSerValThrAsnSerGlnMetSerGlyGluIleProGlyAlaSerProAspHis 168
Db 3046 ATCCCATCGCTCACTCACGACGATCTCAGTGAAATCTCTGGAGCGTCCCTCAT--- 2990
QY 169 HisMetSerProThrGlyAsnIleSerArgArgAla---ProPheProTyrValAsn 187
Db 2989 CATATGATGTTCTCCGTTGGGAACATTGGCAGACGTTGGGCATCCATTTCCCTATGTGAAC 2930
QY 188 HisSerProAsnProSerArgGluPheSerGlySerIleGlyAsnValAlaTrpLysGlu 207
Db 2929 CATTCACCAACCCCATCAAGGAGTTCCTGCTAGCTTGGCAATGTTGTCATGGAAGAG 2870
QY 208 ArgValAspGlyTrpLysMetLysGlnAspLysGlyAlaIleProMetThrAsnGlyThr 227
Db 2869 AGAGTAGATGCTGGAAATGAAG---GATAAAGTGCATTCCTCCATGGCTAATGTACT 2813
QY 228 SerIleAlaProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThrGluTyrAsn 247
Db 2812 AGCATTTCTCTTCAGAAAGTTCGTGGGTTCGTGACATTTGATGTCATCTACTGACTATAAC 2753
QY 248 MetGluAspAlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLysValProIle 267
Db 2752 ATGGAGGACGCTTACTGAATGATGAACACGCAACCTCTATCGAGAAAGTGCACAT 2693
QY 268 AlaSerLysIleAsnProTyrArgMetValIleValLeuArgLeuValLeuSer 287
Db 2692 TCGTCATCCAGATAAATCCGTACAGGATGTTTATTTGTCCTCCGATGATGTCCTATGT 2633
QY 288 IlePheLeuHisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeuTrpLeuLeu 307
```

Db 2632 ATATTCCTGCATACCGTATACAAATCTCTGGCGTAATGTCATACCCGCTATGGTTGCTC 2573
Qy SerValIleCysGluLeuTrpPheAlaLeuSerTrpIleLeuAspGlnPheProLysTrp 327
Db 2572 TCTGTCATATGTGAGATTGGTTGGTTGCTCTATCTCGGATCTCGGATCAGTTCGCCAAGTGG 2513
Qy PheProIleAsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAspArgGluGly 347
Db 2512 TCCCAATCAACCGTGAACACTTACCTGGATAGGCTGGCTTTAAGGTATGACCGGGAAGGT 2453
Qy GluProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspProLeuLysGlu 367
Db 2452 GAACCACTCAATGGCTCTCTGTTGACATTTTGTCTGAGTACTGTGGATCCTATGAAGAAGA 2393
Qy PheProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAsp 387
Db 2392 CCTCCTCTTGTCACTGCCAATACTGTGCTCTCCATCCTTGTCTGGATTACCCCTGTTGAT 2333
Qy LysValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheAspAlaLeuAla 407
Db 2332 AAGGTATCTTGTCTATGTCTGATGATGGAGCTGCAATGTTAACTTTGATGGCTGTGCT 2273
Qy GluThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAspIleGluPro 427
Db 2272 GAGACTTCAGAAATTTGCTAGAAAATGGGTACCGTTCTGTAAAAAGTACACATAGAACCC 2213
Qy ArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGlnPro 447
Db 2212 AGAGCTCCAGAGTGGTACTTTTGCACAGAAAATTTGATTACTTGAAGACAAAGTTCAGGCT 2153
Qy SerPheValLysAspArgAlaMetLysArgGluTyrGluGluPheLysIleArgIle 467
Db 2152 TCTTTTGTAAAGATCGTCGTGCATGACAGAGGAATATGAAGAATTTAAAGTTTCGTGTT 2093
Qy AsnAlaLeuValSerLysAlaLeuLysValProGluGluGlyTrpIleMetGlnAspGly 487
Db 2092 AATGCTCTGTCGCCAAGGACAGAAAGTTCCTGAGAAAGATGGATTATGCAAGATGGC 2033
Qy ThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGly 507
Db 2032 ACACCTTGGCTCGTAAACAATACAGAGACCATCTCGAATGATTACAGGTTTTCTTGGT 1973
Qy HisSerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArg 527
Db 1972 CATAGTGGAGGCTTGATACCGAAGCAATGAGCTTCCTGTTTGTAGTCTATGTGTCGT 1913
Qy GluLysArgProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuValArg 547
Db 1912 GAGAAAGCTCTGGATTCCAAACACCATPAAAAAGCTGGTGTATGAATGCATGTTGCT 1853
Qy ValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIle 567
Db 1852 GTATCAGCTGTCCTTACTAATGGACAATCTTGTGAATCTTGATCTGATGATCCTACATC 1793
Qy AsnAsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyPro 587
Db 1792 AACAAATAGCAAGCTCTCCGAGAGGCTATGTGCTTCCTTATGGATCCAAACCTAGGAAGG 1733
Qy GlnValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArgTyr 607
Db 1732 CGTGTCTGTATGTCCAATTCCTCAGAGGTTTGACGGTATCGATAGAAATGATCGATAT 1673
Qy AlaAsnArgAsnThrValPheAspIleAsnLeuArgGlyLeuAspGlyIleGlnGly 627
Db 1672 GCAACAGGAACACCGTGTGTTTTTCGATATTAATTTGAGAGTCTCGATGGTCTCCAAAGA 1613
Qy ProValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluPro 647
Db 1612 CCAGTTTATGTGGAAACCGTGTGTGTTCAACAGAACAGCTCTTTATGGTTATGAACCC 1553
Qy ProIleLysAlaLysLysProGlyPheLeuAlaSerLeuCysGlyGlyLysLysAla 667

Db 1552 CCCATTAGCAGACAGAGGCCAGGTACTTCTCTCGCTTTGTGGGGACGAAAGAAGACA 1493
Qy SerLysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisValAspSerSer 687
Db 1492 AAAAAGTCCAAAGAGAGACGACGAAAGAAAAGTACACAAACATGTGGACAGTTCT 1433
Qy ValProValPheAsnLeuGluAspIleGluGlyValIleGluGlyAlaGlyPheAspAsp 707
Db 1432 GTGCCAGTTTAACTTGAAGATATAGAGGAAGGATTAAGGTTCTGGATTGATGAT 1373
Qy GluLysSerValLeuMetSerGlnMetSerLeuLysLysArgPheGlyGlnSerAlaAla 727
Db 1372 GAGAAATCACTACTGATGCTCAATGAGCTTACAGAAAAGATTTCGTAATCTAGTGT 1313
Qy PheValAlaSerThrLeuMetGluTyrGlyGlyValProGlnSerSerThrProGluSer 747
Db 1312 TTTGTAGCTCCACTCTGATGGAATATGGTGGTTCCTCAATCCCAACTCCAGAACT 1253
Qy LeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSerGluTrpGly 767
Db 1252 CTTCTGAAGAGGCCATACATGTTATCAGCTGGCTATGAACACAAAAGTACTGGGA 1193
Qy ThrGluIleGlyTrpIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMet 787
Db 1192 ACTGAGATTGGATGGATCTATGGTCTCTGTACAGAAGATATCTCACTGGATTCAAGATG 1133
Qy HisAlaArgGlyTrpArgSerValTyrCysMetProLysArgProAlaPheLysGlySer 807
Db 1132 CATGCACGTGGCTGGCGGTCAATCTACTCATGCCCTAAGCAGCAGCATTTCAAGGGTCT 1073
Qy AlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerVal 827
Db 1072 GCCCAATTAATCTTTCCGATCGTCTGAACAGCTGCTTCGGTGGCTCTTGGTTCTGTG 1013
Qy GluIleLeuPheSerArgHisCysProLeuTrpTyrGlyTyrGlyArgLeuLysPhe 847
Db 1012 GAAATCTTTTACGAGCAATTCCTCATATGTTATGGGTACGAGACGCGCTTAAGTTC 953
Qy LeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeu 867
Db 952 CTGAGAGATTGCTTACATCAACACCACCTATTATTCACCTCACATCGATTCACCTCTC 893
Qy ValTyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIle 887
Db 892 TTGTACTGCATATTCACAGCTATCTGCTCTCTCAGTGGGAAGTTTATCATACAGAGATT 833
Qy SerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIle 907
Db 832 AGCAACTTTGCGAGTATTTGGTTTATCTCTCTCTTCTGTCAATCTTTGTACTGGTATC 773
Qy LeuGluMetArgTrpSerGlyValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrp 927
Db 772 CTTGAGATGAGGTGGATGGTGTGGTATCGATGAATGGTGGAGAAATGACAGTTTGG 713
Qy ValIleGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeu 947
Db 712 GTTATTGGAGTATTTCCGCTCATCTATTTCGGTCTCTCCAGGCTCTGTGTAAGTGCTCT 653
Qy AlaGlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGlyAspPhe 967
Db 652 GCTGGAATTGATACAGCTTCACTGTCCACTCAAAAGGCTCTGATGAAGAAGGCGATTTC 593
Qy AlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIleProProThrThrIleLeuIle 987
Db 592 GCTGAGCTTACATGTTTCAAGTGAGCAACCTCTGTATCCCGACCGACCTATTTTATC 533
Qy IleAsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSer 1007
Db 532 ATCAACTGTGCGGTGTTGTTGCTGATCTCTCCAGCTATCAACAGTGGCTACAGTCA 473
Qy TrpGlyProLeuPheGlyLysLeuPheAlaPheTrpValIleValHisLeuTyrPro 1027
Db 472 TGGGCTCGGCTCTTTGGGAAGCTCTTCTTTCCTTCTGGGTGATGTTGCCACCTGTACCCC 413

```
QY 1028 PheLeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleValTtpAla 1047
Db |||||
412 TTCCTTAAGGGTCTCATGGTGGGAGAACCGTACACCGACCATGTTGTTGTTGGGCC 353
|||
QY 1048 ValLeuLeuAlaSerIlePheSerLeuLeuTtpValArgValAspProPheThrThrArg 1067
Db ::|||
352 ATCCTCCTTGCCATCGATCTTCTCTGCTGCTGGTTCGTATCGATCCATTCACCCCGT 293
|||
QY 1068 LeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db ::|||
292 GTCACAGGCCCTGATACCCCAAAATGTGTATCAACTGC 254
|||
```

Search completed: August 23, 2004, 01:08:01
Job time : 1258 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2004, 00:36:19 ; Search time 29 Seconds

(without alignments)
3582.305 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MDGDADALKSRHGAGDVCO.....VDPFTTLAGPNITQTCGINC 1080

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4507	78.0	1065	2 T52054	cellulose synthase
2	3963.5	68.6	1081	2 T05351	cellulose synthase
3	3801	65.8	1026	2 T51579	cellulose synthase
4	3778.5	65.4	1065	2 F84649	probable cellulose
5	3666.5	63.5	1084	2 T08583	cellulose synthase
6	3583	62.0	1081	2 T52028	cellulose synthase
7	3579	61.9	1088	2 H84604	probable cellulose
8	3477	60.2	974	2 T10797	cellulose synthase
9	3302.5	57.2	958	2 T04870	cellulose synthase
10	2816	48.7	685	2 T10800	cellulose synthase
11	2315	40.1	1181	2 D86157	hypothetical prote
12	2257	39.1	1111	2 T05646	hypothetical prote
13	2220.5	38.4	1145	2 T51546	cellulose synthase
14	2212.5	38.3	1036	2 D84741	probable cellulose
15	2163.5	37.4	979	2 C86446	probable cellulose
16	1941.5	33.6	583	2 T02209	cellulose synthase
17	1094	18.9	757	2 T02561	probable cellulose
18	1076	18.6	712	2 T02552	cellulose synthase
19	1057.5	18.3	755	2 T02553	cellulose synthase
20	1034.5	17.9	748	2 T02560	cellulose synthase
21	983	17.0	689	2 T08918	hypothetical prote
22	963.5	16.7	686	2 T08919	hypothetical prote
23	959.5	16.6	727	2 T08920	hypothetical prote
24	943.5	16.3	828	2 E71417	hypothetical prote
25	872.5	15.1	710	2 B71417	hypothetical prote
26	430	7.4	693	3 AF2275	cellulose synthase
27	289.5	5.0	759	2 D70422	cellulose synthase
28	271.5	4.7	322	2 T12093	TGACG-motif bindin
29	270	4.7	326	2 T08591	TGACG-motif bindin

30	264	4.6	874	2 AB0985	probable polysacch
31	259.5	4.5	1596	2 T31338	cellulose synthase
32	258	4.5	888	2 E91180	probable cellulose
33	257	4.4	692	2 S47754	hypothetical prote
34	257	4.4	888	2 H65151	hypothetical 101.6
35	255.5	4.4	729	2 A98320	cellulose synthase
36	255.5	4.4	729	2 AD2963	cellulose synthase
37	255.5	4.4	861	2 T39714	cellulose synthase
38	251.5	4.4	888	2 G86026	probable cellulose
39	239.5	4.1	322	2 T08592	TGACG-motif-bindin
40	238.5	4.1	754	2 A43735	bcsA protein - Ace
41	232	4.0	664	2 A95889	probable cellulose
42	207	3.6	723	2 C36963	cellulose synthase
43	179	3.1	768	2 B97083	glycosyltransferas
44	164	2.8	654	2 AB2932	beta 1,3 glucan sy
45	164	2.8	654	2 D98350	hypothetical prote

ALIGNMENTS

RESULT 1

T52054

cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000

C:Accession: T52054

R:Arioli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.; H

A:Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.

A:Reference number: Z13745; MUID:98111412; PMID:9445479

A:Accession: T52054

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1065 <ARI>

A:Cross-references: EMBL:AF027174; PIDN:AAC39336.1

C:Genetics:

A:Note: Ath-B

C:Function:

A:Description: EC 2.4.1.-; cellulose synthase [validated, MUID:98111412]; involved in as

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 78.0%; Score 4507; DB 2; Length 1065;

Best Local Similarity 78.4%; Pred. No. 0;

Matches 836; Conservative 98; Mismatches 107; Indels 26; Gaps 8;

Qy 19 CQICADGLGTTLDGVDFTACDVCRCPCYERKEGTCQACLCQCKTKYKRRHSGPAIR 78

Db 20 CQICSDNVGKTVGDGRFVACDICSFPVCRPCYERKDGNGQSCPCQCKTRYKRLKSGPAIP 79

Qy 79 GEEGDDTDADGS-DFNYPASGTDQKKTADNRSMWMTGGSGNVGHPKYSDEIGLS 137

Db 80 GDKDEGLADEGTVFEFNYP-----QKEKTSERMLGWHLTRGKGEMGEPOYDK----- 127

Qy 138 KYDSGEIPRGVPSVINSQ-MSGBIPGASPDHMHMPTGNISRRAPPY-VNHSNPSRS 195

Db 128 -----EVSHNHLPLKLTQRDTSGEFSAASPERLSVSSTIAGGKRLPYSSDVNQSPNRIV 182

Qy 196 FSGSIGNVAMKERYVDGWMKQDKGAIPTMTGTSTIAPSEGRAATIDIDASTEYNMEDALLND 255

Db 183 DPVGLGNVAMKERYVDGWMKQDKGAIPTMTGTSTIAPSEGRAATIDIDASTEYNMEDALLND 238

Qy 256 ETROPLSRKVPDASSKINPRMIVLRLVLSLFLHRLTNPNVNPVPLMLLSVICIWF 315

Db 239 EAROPLSRKVPDASSKINPRMIVLRLVLSLFLHRLTNPNVNPVPLMLLSVICIWF 298

Qy 316 ALSWLDQFPKVPINRETYLDRALRYDREGESQLAAVDIFVSTVDLKEPPVIVTANT 375

Db 299 ALSWLDQFPKVPINRETYLDRALRYDREGESQLAAVDIFVSTVDLKEPPVIVTANT 358

Qy 376 VLSILAVDYPVDKVCYVSDGASMLTFDALAETSEFARKWVPFKKYDIEPRAPEFYFC 435

Db 359 VLSILAVDYPVDKVCYVSDGASMLTFDALAETSEFARKWVPFKKYDIEPRAPEFYFC 418

QY	436	OKIDYLDKQVQPSFVKDRAMREYEEFKIRINALVSKALKVPEEGWIMQDGTTPPGNNT	495
Db	419	AKIDYLDKQVQPSFVKDRAMREYEEFKIRINALVSKALKVPEEGWIMQDGTTPPGNNT	478
QY	496	RDPHGMIOVFLGSHGGLDTEGNEPLRLVVSREKPGFGQHHKKAGAMNALVRVSAVLITNG	555
Db	479	GDHPGMIQVFLGSGGLDAEGNELPRLVVSREKPGFGQHHKKAGAMNALVRVSAVLITNG	538
QY	556	QWMLMLDCDHYINNKAVREACFLMDPNLGRQVQCVVQPPQDFDGLDRNDRYANRNTVFF	615
Db	539	PFILMLDCDHYINNKALREACFLMDPNLGRQVQCVVQPPQDFDGLDKNDRYANRNTVFF	598
QY	616	DINLRGLDGIQGPVVVGTGCVFNRTAIVGYEPIKA--KKPGFLASLCGGKKKASKSKR	673
Db	599	DINLRGLDGIQGPVVVGTGCVFNRTAIVGYEPIKVKVHKPKSLLSKLGGSRKKSKAKK	658
QY	674	SSDKKSKNKHVDSPVFVFNLEDEEGVEGAGFDDKSKVLMQMSLEKRFQSGAAVFASTL	733
Db	659	ESDKKSKGRHTDSTVPFVFNLDIEEGVEGAGFDDKALLMSQMSLEKRFQSGAVFASTL	718
QY	734	MEYGVQPSSTPESLLKEAHIHVSICGYEDKSWGTETIHWIYGSVTEDILTGFKMARGWR	793
Db	719	MENGVPVPSATPENLLKEAHIHVSICGYEDKSWGMEIGWIYGSVTEDILTGFKMARGWR	778
QY	794	SVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVELFSRHCPLMVGYGGRLLKELERPAY	853
Db	779	SYICMKPLPAKGSAPINLSDRNLQVLRWALGSVELFSRHCPIMWYNGRLKFLERPAY	838
QY	854	INTTIYPLTSLPLLYCYILPAICLTGKFMPEISNLASIFIALFLSIFATGILEMRWS	913
Db	839	VNTTIYPTITSIELLMYCTLLAVCLFTNQFIIPQISNIASIFLFLSIFATGILEMRWS	898
QY	914	GVGIDEMWRNEQFWIGGISAHLPFAVQGLLKVLAGIDTNFTVTSKANDEEGDPAELIYMF	973
Db	899	GVGIDEMWRNEQFWIGGVSAHLFAVQGLILKVLAGIDTNFTVTSKASDEGDPAELIYLF	958
QY	974	KWTTLLIPTTILLINMGVWAGTSVAINSGVQSGWGLFGKLFPAFWIIVHLIYPLFKGLM	1033
Db	959	KWTTLLIPTTILLIIVNLGVVAGVSAVNSGVQSGWGLFGKLFPAFWIIVHLIYPLFKGLM	1018
QY	1034	GRQNRTPITIVWVLLASIFSLMLWVRDPTFTRLAGPNIQTCGNC	1080
Db	1019	GRQNRTPITIVVWSILLASIFSLMLWVRIDPFTSRVTGPDILECGNC	1065
RESULT 2			
T05351			
cellulose synthase (EC 2.4.1.-) catalytic chain RSW1 - Arabidopsis thaliana			
N:Alternate names: protein F8B4.110			
C:Species: Arabidopsis thaliana (mouse-ear cress)			
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999			
C:Accession: T05351			
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De			
ewes, H.W.; Mayer, K.F.X.; Schueller, C.			
submitted to the Protein Sequence Database, February 1999			
A:Reference number: Z15409			
A:Accession: T05351			
A:Molecule type: DNA			
A:Residues: 1-1081 <BEV>			
A:Cross-references: EMBL:AL034567			
A:Experimental source: cultivar Columbia; BAC clone F8B4			
C:Genetics:			
A:Gene: RSW1			
A:Map position: 4			
A:Intron: 27/3; 93/1; 150/3; 189/1; 252/2; 341/2; 456/3; 502/3; 544/3; 615/3; 704/1; 76			
C:Keywords: glycosyltransferase; hexosyltransferase; P-loop			
Query Match 68.6%; Score 3963.5; DB 2; Length 1081;			
Best Local Similarity 68.0%; Pred. No. 3.2e-282;			
Matches 733; Conservative 138; Mismatches 168; Indels 39; Gaps 15;			

QY	2	DGDADALKSRRGAGDVCOICADGLGTTLDGDFVACDVCVRFPVPCPYEHERKEGTQAC	61
Db	25	DGDTPLKNNM---QIQICQICGDDVGLAETGDFVACNECAPVCPVPCPYEYERKDTQCC	81
QY	62	LOCKTKYKRRHSPAIRGEEGDDTDADGSDTFNYFASGTEDQOKIADRMRSWRMTGGS	121
Db	82	POCKTRFRHRHSRPRVEGDEDDDDIENEENYAQGANKARHQRHGEFSS-----	133
QY	122	GNVGHPKYDSGIGLSKYD---SGEIPRGYVPSVINSQMSGELPGASPDHMMKSPGNTS	178
Db	134	-----SSRHESQIPILLTHGHTVSGEIRTPDTSVTRT---SGPL-SPSDRNAISSP--YID	184
QY	179	RRAPPYVNHSPNPSREF--SGSIGNAVWAKERVDGWMKQDKGAIPTWNGTSTIAPSGRAA	237
Db	185	PRQPVV--RIVDPKOLNSYGLGNVWKERVEGWKLKQEKMLQMTG---KYHEGKGG	238
QY	238	TDIDASTEYNEBDALINDETQPLSRKVPKVIASKINPYRMVILVRLVLSIFLHVLNTP	297
Db	239	-EIE-GTSGNGBELQMADDTRLPMSRVVPIPSRLTFYRVVILRLIILCLFQLVTRTHP	296
QY	298	VENAYPLMLLSVICETWEALSWILOPKWFPINRETYLDRLALRYDREGSPSOLAADI	357
Db	297	VKNAYPLMLTSVCEIWFAPSWLLDQFPKWYPINRETYLDRLAIRYDRDGSPSOLVFDV	356
QY	358	FVSTVDPLEKPEPPIVTANTVLSILAVDYPVDKVCVSDDGASMLTFDALAETSEFARKW	417
Db	357	FVSTVDPLEKPEPPIVTANTVLSILSDVDYPVDKVCVSDDGASMLTFSEISETAEFAKKW	416
QY	418	PFVKYVDIEPRAPBYFCQIDYLDKQVQPSFVKDRAMREYEEFKIRINALVSKALKV	477
Db	417	PFCKXFNIEPRAPBYEYFAQIDYLDKQIQPSVFKERRAMREYEEFKIRINALVAKAQI	476
QY	478	PEEGWIMQDGTTPGNNTRDHPGMLOVFLGSHGGLDTEGNEPLRLVVSREKPGQHHK	537
Db	477	PEEGWIMQDGTTPGNNTRDHPGMIQVFLGSHGGLDTEGNEPLRLIIVSREKPGQHHK	536
QY	538	KAGAMNALVRVSAVLITNGQYMLNLCDDHYINNKAVREACFLMDPNLGPQVQVFPQR	597
Db	537	KAGAMNALIRVSAVLITNGAYLLAVDCDHYFNNSKAIKEAMCFMMDPAIGKKCCYVFPQR	596
QY	598	FGIDRNDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIVGYEPPKAK--KFG	655
Db	597	FGIDILHNDRYANRNTVFFDINMKGLDGIQGPVYVGTGCVFNRTAIVGYEPPKAK--KFG	656
QY	656	FLA-SICGKKKASKKSSDKKSNKHVDSPVFVFNLEDEEGVEGAGFDDKSKVLS	714
Db	657	IIVKSCGSRKKGKSKKTYNEKRRINRSDNAPLFNMEDIDEGF--GYDDERSILMS	714
QY	715	QMSLEKRFQGSAAVFASTIMEYGVQPSSTPESLLKEAHIHVSICGYEDKSEWGTEIGTY	774
Db	715	QRSVEKRFQGSVPFIAATFMEQGGIPPTNPATLLKEAHIHVSICGYEDKTEWGKEIGWIY	774
QY	775	GSVTEDILTGFQKHARGWRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRH	834
Db	775	GSVTEDILTGFQKHARGWISYICNPRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRH	834
QY	835	CPLWYGYGRGLFLERFAYINTTIYPLTSLPLLYCYILPAICLTGKFMPEISNLASTW	894
Db	835	CPWIYGYGRGLRLERAYINTVIPTSIPUIACILPAFCITDTRFIPIPSINVASIW	894
QY	895	FIALFLSIPATGILEMRWSGVGIDEMWRNEQFWIGGISAHLPFAVQGLLKVLAGIDTNF	954
Db	895	FILLFISIAVTGILELRWSGVSIEDWRNEQFWIGGTSAHLFAVQGLLKVLAGIDTNF	954
QY	955	TWTSKANDEEGDPAELYMFKWTLLIPTTILLIINMGVWAGTSVAINSGVQSGWGLFGK	1014
Db	955	TWTSKATDEGDPAELYIFKWTALLIPTTILLIINMGVWAGTSVAINSGVQSGWGLFGK	1014
QY	1015	LEFPAFWIIVHLIYPLFKGLMGRQNRPTIVVWVALLASIFSLMLWVRDPTFTRLAGPN	1072
Db	1015	LEFALWVIAHLIYPLFKGLLGRQNRPTIVVWVALLASIFSLMLWVRINFPVD--ANPN	1070


```

Db      824 LRWALGSVEIFLSRHCHCPIWYGGGLKWLRFESVINSVVPWISLPLIVYCSLPVAVCLLT 883
Qy      880 GKTIMPISNLASITWIALFSLFATGILEMRSGVIGIDEWRNEQWVITGGISLAHLPV 939
Db      884 GKFTVPEISNAYAGILFIMFISIAVTGILEBQWGVGIDWWRNEQWVITGGASHLFAL 943
Qy      940 FOGLLKVLAGIDNFTVTSKANDEEGFAELYFKEWTTLLIPPTILLINMGVVACTSY 999
Db      944 FOGLLKVLAGVNTFTVTSKAAD-DGAFSELYIFKWTLLIPPTILLINMGVIGVSD 1002
Qy      1000 AINSQYSGWGLFGLKFAFVWVHLVFLKGLMGRQNRTPPTIVWAVLLASIFSLWV 1059
Db      1003 AINSQYDSWGLFGLRFLFALWVHLVFLKGLMGKQDKMPTIIVVMSILLASITLLWV 1062
Qy      1060 RVDPTFTRLAGPNIQTGCINC 1080
Db      1063 RVNPFVAK-GGPVLEICGLNC 1082

RESULT 6
T52028
cellulose synthase [imported] - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C;Accession: T52028
submitted to the EMBL Data Library, May 1998
A;Reference number: Z25890
A;Accession: T52028
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1081 <JOS>
A;Cross-references: EMBL:AF062485; PIDN:AAC29067.1

Query Match      62.0%; Score 3583; DB 2; Length 1081;
Best Local Similarity 61.1%; Pred. No. 2.7e-254;
Matches 676; Conservative 162; Mismatches 198; Indels 72; Gaps 15;

Qy      1 MDGADAA-LKSGRHGAGDVQCI CADGLGTLDGVFTACDVCRFPVCRPCYEHKRGQTQ 59
Db      17 INADENAIRSVQSLSGQTQICRDEITELTVGDEPFVACNECAPVCRPCYEVERREGNQ 76
Qy      60 ACLOCKTKYKRRHSGPAIRGEGDDTDADGSDNPYPASGTE-DQKQIADRMRSWRMT 118
Db      77 ACPOCKTRFRRLKSGPRVGEDEDDDDLDLNEFYGNNGIGFDQ---VSEGMSISRNS 133
Qy      119 GGSNGVGHPRKYDSGEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHH----- 169
Db      134 -----GFPQSD-----LDSAP-FGSQIPLITYGDEDEVEI---SSDRHALVPPSLG 175
Qy      170 -----MMSTGNISRRAPPYVNHSPNRPFGSGISGNVAKERVVDGWMKQDKG 219
Db      176 GHGNRVHPVSLSDPTVAHRRLMVQPKDLAVY-----GYGSVAMKDRMBEWRKQNE- 227
Qy      220 AIPWNTGTSIAPSGRAATIDIDASTEYNMEDALLNDETROPLSKVPITASSKINPYRMVI 279
Db      228 -----KLQVRHGGDDPDFEGDDADPFM-----MDEGRPLSMKPIPKSKINPYRMVI 276
Qy      280 VLRVLVLSIFLHYRLTNPVRNAYPLWLLSVICETWFALSWILDQPKWFPINRETYLDRL 339
Db      277 VLRVLVILGLPHYRLHPVDAYALWLSVICEWFAVSVWLDQFPKWPYIERETYLDR 336
Qy      340 ALRYDRGERSQLAANDIFVSTVDPLKEPPIVTANTVLSILAVDPVDKVCYVSDDGAS 399
Db      337 SLRYEKGKPSGLSPVDVFNFTVDPLKEPPLITANTVLSILAVDPVDKVCYVSDDGAA 396
Qy      400 MLTFDALAETSEFAKRWPFVKKYDIEPRAPEFFECOKIDYLDKQVPSFVKDRAMKRE 459
Db      397 MLTFEALSETAEFAKRWPFVKKYDIEPRAPEFFECOKIDYLDKQVPSFVKDRAMKRD 456
Qy      460 YEEFKIRINALVSKALVPBEGTWQDGTWPNGNTRDHPGMIQVFLGSHGGLDTEGNEL 519

```

```

Db      457 YEEFKVKINALVATAQKVPEDGWTQDGTWPGNSVRDHPGMIQVFLGSDGVRDVNNEL 516
Qy      520 PRLVYYSREYSPGQHHKKAGAMNALVRYSAVLTNGOYMLNLDCHYINNSKAVREAMCF 579
Db      517 PRLVYYSREKPGFDHKKAGAMNSLRVSGVLSNAPYLLNVDCDHYINNSKALREAMCF 576
Qy      580 LMDPNLGPQVYQVQFPORFDGIDRNDRYANRNTVFFDINLRGLDIOGQPVVVGTCVFNRR 639
Db      577 MMDPOSKKICYQVQFPORFDGIDRHDYRNRVVFFDINMKGLDGLGPIYVGTGCVFRR 636
Qy      640 TAIYGYEPPIKAKKPGFLAS-----LCGKKKASKSKRRSSDKKSKNKHVDSVSVFN 692
Db      637 QALYGFDPAPKKKGGPRKTCNCPKWCCLLCFGSRKNRKAFTVAADKKKKNR--EASKQIHA 694
Qy      693 LEDIEEGVAGFDDEKSVLMSQMSLEKRFQGSNAFVASTILMEYGGVQPSQSTPESILKEA 752
Db      695 LENIEGRGHKVLNVBOSTEAMQMKQKQYGSQFVFVASARLENGGWARNASPACLLKEA 754
Qy      753 IHVISCYEDKSEWGTETIGWYGSVTEDILTDFGKHARGWRSVYVCMPEKRFAPKGSAPINL 812
Db      755 IQVISRGYEDKTEWKEIGWYGSVTEDILTGSQSHSGRHHVYCTPKLAAPKGSAPINL 814
Qy      813 SDRNLQVLRWALGSVEILFGRHCPMYGYGGRUKFLERPAYINTIYPLTSLPLLVYCIL 872
Db      815 SDRLQVLRWALGSVEILFGRHCPMYGYGGRUKFLERPAYINTIYPLTSLPLLVYCSL 874
Qy      873 PAICLLTGKEMPEISNLASIFALFSLFATGILEMRSGVIGIDEWRNEQWVITGGI 932
Db      875 PAICLLTGKFIPEISNAYSILFMAUFGSSIAITGILEMQMGKVIDDWMWRNEQWVITGGV 934
Qy      933 SAHLFAVFOGLLKVLAGIDNFTVTSKANDEEGFAELYFKEWTTLLIPPTILLINMGV 992
Db      935 SAHLFALFOGLLKVLAGIDNFTVTSKAAD-DEEFSDLYLFKWTSLIIPMTLLIINVG 993
Qy      993 VVAGTSYAINSGYQSWGPFGLKFFAFWVIVHLYPFLKGLMGRQNRTPPTIVWAVLLAS 1052
Db      994 VIVGVSDAISNGYDSWGLFGLRFLFALWVHLVFLKGLMGKQDKMPTIIVVMSILLAS 1053
Qy      1053 IFSLLWVRVDPFTTRLAGPNIQTGCINC 1080
Db      1054 IFTLLWVRVNPVFAK-GGPVLEICGLDC 1080

RESULT 7
H84604
probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: H84604
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vauken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84604
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1088 <STO>
A;Cross-references: GB:AE002093; NID:g4417271; PIDN:AAD20396.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g21770
A;Map position: 2

Query Match      61.9%; Score 3579; DB 2; Length 1088;
Best Local Similarity 61.1%; Pred. No. 5.3e-254;
Matches 677; Conservative 156; Mismatches 199; Indels 76; Gaps 20;

Qy      4 DADALKSGRHGAGDVQCI CADGLGTLDGVFTACDVCRFPVCRPCYEHKRGQTQACLO 63
Db      24 DTAIRSABELSGQTKICRDEITELTDNGEPFFACNECAPFTCRPCYEVERREGNQACPQ 83
Qy      64 CKTKYKRRHSGPAIRGEGDDTDADQ-----GSDFNYPASGTEQDKQIADRMRSWRN 117

```

```
Db      84  CQTRYKRIKSPRVEGDEDD--DIDDLEHEFGMD---PEHVTE-----AALYVMRLN 132
QY      118  TG-GSGNVCHPKYDSEIGLSKYDSEIGPRGVVPSVTNSQMSGEIPGASPDHMMSP---- 173
Db      133  TGRGTDEVSH-----LYSAS--PGSRVPLLTYYCDESDM--YSDRHALVPPST 177
QY      174  -TGNLSRRAPF-----PVNHSNPNPREFS--GSIGNVAKERVDMQWKO--DKGA 220
Db      178  GLGNRVHVPFDTSPASITHRPV-----PQDLTVYGYGSVANKORMEVWKKQIEKIQ 232
QY      221  IPTMNGTSTIAPSEGRAATDIDASTEYNMEDAL--LNDETROPLSRKVPIASSKINPYMVI 279
Db      233  VVKNERNVNDGCGDGFVDELD-----DPGLPMWDEGRQPLSRKLPTRSSRNINPYMLI 285
QY      280  VLRLVLSLFLHYRLTNVRNAYPLWLLSVCEIWFALSWILDDPKWPPINRETYLDR 339
Db      286  FCRLLAILGLFFHYRLHPNDAGLWLTSVCEIWFASWILDDPKWPIERTYLYDR 345
QY      340  ALRYDREGEPSQAAAVDIFVSTVDPLKEPPIVTANTVLSILAVDYPVDKVSQVSDDGAS 399
Db      346  SLRYEKEGKPSLAPVDVFSVTDPLKEPPLITANTVLSILAVDYPVEKVACVSDGAA 405
QY      400  MLTTPALAESEFARKWVPFVKXYDIEPRAPPEFCQIDYLDKQVQPSFVKDRRAMKRE 459
Db      406  MLTFEALSYTAEFARKWVPFCKKFSIEPRAPPEWFSQKMDYLKHKVDPDAFVMERRAMKRD 465
QY      460  YEEFKIRINALVSKALVPEEGWIMODGTPWPGNTRDHPGMIOVFLGHSGGLDTEGNE 519
Db      466  YEEFKVKNALVSVSQKPEDGWTQDGTWPFGNVRDHPGMIOVFLGHSGVCDMDGNE 525
QY      520  PRLVVSREKRGPFQHHKKAGAMNALVRVSAVLINGQYMLNDCDHYINNSKAVREAMCF 579
Db      526  PRLVVSREKRGPFQHHKKAGAMNSLIRVSAVLSNAPYLLNVDCDHYINNSKAIREAMCF 585
QY      580  LMDPNLGVQVCVQPPQDPDIDRNDRYANRNTVFFDINLRGLDGIQGVVYVGTGVNRR 639
Db      586  MMDPQSGKKICVQPPQDPDIDRDRYSNRNVFFDINMKGLDGIQGPVYVGTGVNRR 645
QY      640  TAIYGEYPIKAKPGFLAS-----LCGGKKKASKKRSDDKKSKNKHVDSVPVFN 692
Db      646  QALYGFDAKKQPPGRTCNWPKWCCCLCCGRRKKTKGVQDNQRRKKP--ETSKQIHA 702
QY      693  LEDIEGVEGAGFDDEKSVLMSQMSLEKRFQOSAAFAVASTLMEYGGVQPSSTPESILLKEA 752
Db      703  LEHIEEGVQVTAENNSSET--AQLKLEKFGQSPVLVASTLLNGGVSNVNPASLLRES 760
QY      753  IHVISCGHEDKSEWGTETIGWYGSVTEIDLITGFKMHARGWSVYCMKRPAPKGSAPINL 812
Db      761  IQVISCGEYEETEWGKEIGWYGSVTEIDLITGFKMHCHGWSVYCMKRAAPKGSAPINL 820
QY      813  SDRLNOVLRAWLGSVEILFSRCHPLWYGGGRKFLERFAYINTTIYPLTSLPLAVYCIL 872
Db      821  SDRLHQVLRAWLGSVEILFSRCHPLWYGGGLKWLERFSYNSVYVYPTSLPLVYCSL 880
QY      873  PAICLLTGKFIPEISNLASIFIALFSLFATGILEMRWSGVGIDEMWRNEQFWVIGI 932
Db      881  PAICLLTGKFIPEISNAGILFLAMFMSIAVTGILEMQWKGIGIDDWRNEQFWVIGV 940
QY      933  SAHLFAVFOGLLKVLAGIDTNPVTISKANDEGDFAEILYMKFWTLLIPPTTILINWVG 992
Db      941  SSHLFAVFOGLLKVLAGYSTNPVTISKAA--DGEFSELYIFKWTSLIPPTTILINI 999
QY      993  VVAGTSYAINSGYQSGWPLFGKLFGLFAFVVIHLYPLKGLMGQRNRTPTIVIVMAVLAS 1052
Db      1000  VIVGSDAINNGYDSGWPLFGELFALWVIVHLYPLKGLKGQRVVTIILVWSILLAS 1059
QY      1053  IFSLLWVRVDPPTTLRAGNIQTGCINC 1080
Db      1060  ILTLLWVRVNPVFSK--DGPVLEICGLDC 1086
```

```
Tl0797
cellulose synthase (EC 2.4.1.-) catalytic chain celA1 - upland cotton
C,Species: Gossypium hirsutum (upland cotton)
C,Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C,Accession: Tl0797
R,Pear, J.R.; Kawagoe, Y.; Schreckengost, W.E.; Delmer, D.P.; Stalker, D.M.
Proc. Natl. Acad. Sci. U.S.A. 93, 12637-12642, 1996
A,Title: Higher plants contain homologs of the bacterial celA genes encoding the catalytic
A,Reference number: Z17152; MUID:97057296; PMID:8901635
A,Accession: Tl0797
A,Status: preliminary; translated from GB/EMBL/DBJ
A,Molecule type: mRNA
A,Residues: 1-974 <PE>
A,Cross-references: EMBL:U58283; NID:gl706955; PIDN:AAB37766.1; PID:gl706956
A,Experimental source: strain Acala SJ-2; fiber
C,Genetics:
A,Gene: celA1
C,Function:
A,Description: involved in the synthesis of cellulose
C,Keywords: glycosyltransferase; hexosyltransferase

Query Match      60.2%; Score 3477; DB 2; Length 974;
Best Local Similarity 61.5%; Pred. NO. 1.4e-246;
Matches 657; Conservative 129; Mismatches 176; Indels 106; Gaps 15;

QY      18  VQCIACADGLGTTLDGDVFTACDVCRPCYHERKEGTQACLOCKTKYKRHRGSPAI 77
Db      8  VCHTCGEHVGLNVNGEPFVACHCECNFPICKSCFFDLKEGRKACLR-----GSPY- 58
QY      78  RGEEDDTDADGSDFNYPASGTEDQKQIADRMRSWRMTTGGSGNVG--HPKYDSEIGL 136
Db      59  -----DENLDD-----VERATGQSTMAA-----HLNKSQDVGIHARHIS---SV 96
QY      137  SKYDSEIPIRGVPSVTNSQMSGEIPGASPDHMMSPGINSRRAPFPVNVHSPSPREF 196
Db      97  STLDS-----EM 103
QY      197  SGSIGNAWKERVQGMKQDKGAIPMTNGTSTIAPSEGRAATDIDASTEYNMEDALLNDE 256
Db      104  AEDNGNSTWKNRVESWKEKKKKKPAT-----KVERAEIPEQOMEDPAPDA 154
QY      257  TRQPLSRKVPPIASSKINPYRMVILRLVVLVLSIFLHYRLTNVRNAYPLWLLSVCEI 316
Db      155  S-QPLSTIIPKSLAPYRTVIIMRLIILGLFHYRTNPVDSAFGLWLTSVCEI 213
QY      317  LSWILDQPKWPPINRETYLDRALRYDREGEPSQAAVDFVSTVDPLKEPPIYANTV 376
Db      214  FSWILDQPKWPPINRETYLDRLSARYEREGERDELAADVDFVSTVDPLKEPPLITANTV 273
QY      377  LSILAVDYPVDKVSQVSDDGASMLTFDALAETSEFARKWVPFVKYKDIETPRAPPEFCQ 436
Db      274  LSILALDYPVDKVSQVSDDGASMLTFESLVETADFARKWVPFCKKFSIEPRAPPEFSQ 333
QY      437  KIDYLDKQVPSFVKDRRAMKREYEEFKIRINALVSKALKVPERGWINODGTPWPGNTR 496
Db      334  KIDYLDKQVPSFVKERRAMKRDYEEYKIRINALVAKAQTDPDEGWTQDGTSGWPGNPR 393
QY      497  DHPGMIOVFLGHSGGLDTEGNEPLRLVVSREKPGFOHHKKAGAMNALVRVSAVLN 556
Db      394  DHPGMIOVFLGYSGARDIEGNEPLRLVVSREKRPYQHHKKAGAMNALVRVSAVLNAP 453
QY      557  YMLNLDCDHYINNSKAVREAMCFMLMDPNLGPQVCVVPQFPQDFGIDRNDRYANRNTV 616
Db      454  FILNLDCDHYINNSKAVREAMCFMLMDPOVGRDVCVVPQFPQDFGIDRNDRYANRNTV 513
QY      617  INLRGLDGIQPVVYVGTGVNRTAIYEPPIKAKPGFLASLGGKKKASKSKRRSSD 676
Db      514  VNMKGLDGIQPVVYVGTGVNRTAIYEPPIKAKPGFLASLGGKKKASKSKRRSSD 573
QY      677  KKSKNKHVDSVPVFNLEIDEEGVAGFD--EKSVLMSQMSLEKRFQSGAAFAVASTIME 735
Db      574  LYRAKREELDAALFNLRID-----NYDEYERSMILSQTSEFKTGLGSSVFIESTIME 627
```

Qy 736 YGVVPQSSPTESLLKEAIHWISCYEDKSWGTEIGWYGSVTEDILTQPMHARGWRSV 795
Db 628 NGGVAESANFSTLLKEAIHWISCYEKTANGKEIGWYGSVTEDILTQPMHCRGWSRI 687
Qy 796 YCMKPRAPAFKGSAPINLSDRINQVRLWALGSVEILFSRHCPFLWVGY-GGRLKFLERFAYI 854
Db 688 YCMFLRPAFKGSAPINLSDRLHQVRLWALGSVEILFSRHCPFLWVGFGGRLKWLQRLAYI 747
Qy 855 NTTIYPTSLTPLLVCYLCPAI CLLTGKFMPEISNLSIWFIAFLFSIFATGILEMRWSG 914
Db 748 NTIYVPTSLFLAYCSPAI CLLTGKFIPTIISNLSVFLGLFSLIIVTAVLELRWSG 807
Qy 915 VGIDWWRNEQFWWIGIGISAHLFAVFOGLKVLGADITNFTVTSKANDEGDFAEALYMPK 974
Db 808 VSIEDLWRNEQFWWIGVSAHLFAVFOGLKMLAGIDITNFTVAKAAD-DADFGELYIVK 866
Qy 975 WTTLLIIPPTILLINMGVVAGTSYAINSGYSQSGPLFGKLFPAFWIVVHLYPFLKGLMG 1034
Db 867 WTTLLIIPPTILLINMGVVAGTSYAINSGYSQSGPLFGKLFPAFWIVVHLYPFLKGLMG 926
Qy 1035 QRNRTPTIVTVAALLASIFSLWVRVDPFTRLAGENI-QTC-GINC 1080
Db 927 QRNRTPTIVVLSVLLASVLSLWVRINPFSVSTADSTTVSQSCLSIDC 974

RESULT 9
T04870
cellulose synthase (EC 2.4.1.-) catalytic chain F28A21.190 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C;Accession: T04870
R;Bavan, M.; Mueller, M.W.; Mueundlein, A.; Felber, R.; Bancroft, I.; Mewes, H.W.; Mayer, submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15387
A;Accession: T04870
A;Molecule type: DNA
A;Residues: 1-958 <BEV>
A;Cross-reference: EMBL:AL035526
A;Experimental source: cultivar Columbia; BAC clone F28A21
C;Genetics:
A;Map position: 4
A;Introns: 60/1; 76/3; 94/1; 149/2; 240/2; 355/3; 401/3; 443/3; 514/3; 662/3; 780/3; 791/3
A;Note: F28A21.190
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match	57.28;	Score	3302.5;	DB	2;	Length	958;
Best local Similarity	59.5%;	Pred.	No. 8.5e-234;				
Matches	638;	Conservative	125;	Mismatches	174;	Indels	135;
Gaps	17;						
Qy	18	VQICADGLGTTLLDGDVFTACDVCRFPVCYEHKEGTQACLOCKTKYKXHRGSPAI	77				
Db	8	ICNTGGEIYGKSGNEFFVACHCSFPICACUEYEFKEGGRICLCRGNPY-----	58				
Qy	78	RGEEGDITDADGSDFNYPASGTEDQKQIADRMRSWRMNTGSGNVGHGPKYDSGEIGLS	137				
Db	59	-----DENVFDDVE-----TKTSKTQSID-----SGI--HARHIS---TVS	89				
Qy	138	KYDSGETPRGVVPVSTNSQMSGEITGASPDHMMMSPTGNISRRAPFPVNVHSPNSPREFS	197				
Db	90	TIDS-----ELN	96				
Qy	198	GSIGNVAKERVDCGKMKQDKGAIPMTNGTISIAPSEGRAATDIDASTEYNMEDALLNDET	257				
Db	97	DEYGNPIWKNVESWKOKDKSKKKKKDKPKATKAEQHEA---QIPTQGHMEDPPPPNTES	153				
Qy	258	--ROPLSRKVPITASSKINPYRMVITVLRLVLSIFLHYRLTNPNVNAVPLMLLSYICEIWF	315				
Db	154	GATDVLVSWVPIPTKTSYRIVILMRLIILALFNRYRITHPVDSAGVLMVTSYICEIWF	213				
Qy	316	ALSWILQFPKWFPIINRETYIDRLALRYDREGESQIAAVDIFVSTVDVDPKKEPPIVANT	375				
Db	214	AVSWLQFPKWSVINRETYIDRLSASFEREGESQIAAVDFFVSTVDVDPKKEPPIVANT	273				

Query 376 VLSILAVDPVDKVCYSVSDGASMLTTFDALAETSEFAKWKVPFVKYKDYIDIEPRAPEYFC 435
Db 274 VLSILALDYPVDKVCYSVSDGAAALSFESLSETADFAKWKVPFCKKYSIEPRAPEYFS 333
Query 436 QKIDYLDKVKQPSFVKDPRAMKREYEEFKIRINALYSKALKVPEEGWIMODGTQWPGNNT 495
Db 334 LKIDYLDKVKQPSFVKERRAMKRDYEEFKIRIMALVAKAQKTEEGWIMODGTQWPGNNT 393
Query 496 RDHPGMIQVFLGHSGLDTEGNELPRLVVVSREKRPFGQHHKKAGAMNALVRVSAVLTNG 555
Db 394 RDHPGMIQVFLGSGARDIEGNELPRLVVVSREKRPFGYQHHKKAGAENALVRVSAVLTNA 453
Query 556 QYMLMLDCHYINNSKAVREAMCFMLMDPNLGPQVCYVQFPQREFGIDRNDRYANRNTVFF 615
Db 454 PFILMLDCHYVNNSKAVREAMCFMLMDPVVQGVQVCFVQFPQREFGIDKSDRYANRNTVFF 513
Query 616 DINLRGLDGIQGPVYVGTGCVFNRTALYGEVPPDKAKKPGFL-----ASLCGGKKKASKS 670
Db 514 DVNMHGLDGIQGPVYVGTGTVFRQALYGYSP-----SKPILPQSSSSSSCCCLTKK--KQ 568
Query 671 KKRSDKKKSNKHVDSPVFNLEDEEGVEGAGFDD-EKSVLMQMSLEKRFQCSAAFV 729
Db 569 PQDPSEIYKDAKBEELDAAIFNLGDL-----NYDEYDRSMLISQTSFEKTFGLSTVFI 622
Query 730 ASTLMYGVGPOSSPESLKLKAHIVISCGYEDKSEWGTGIEGWYGSVTDILTFGPKMHA 789
Db 623 ESTLMENGVPDPSVNPSTLKEAHIVISCGYEBKTEWGEIGWYGSITDILTFGKMHCH 682
Query 790 RGWRSVYCMKPKPAKPGSAPINLSORLNQVLRWALGSVEILFSRHCPCLWYG-YGGRUKFL 848
Db 683 RGWRSIYCMURPAKPGSAPINLSORLHQVLRWALGSVEILFSRHCPCLWYGCGRUKLL 742
Query 849 ERFAVINTTIVPLTSLPLVYCIPLAICLLTGKFIPEISNLSAIFIALFLSIFATGIL 908
Db 743 QRLAYINTIVPFTSLPLVAYCTPLPAICLLTGKFIPTLSNLASMLFL----- 790
Query 909 EMRWGSGVIDEWRNEQFVWIGISAHLPFAVFGLLKVLAGIDTNTFTVTSKANDEBGDFA 968
Db 791 -----GVSTDELWRNEQFVWIGVSAHLPFAVFGFLKMLAGLDTNFTVTSKTADDL-EFG 844
Query 969 ELYMFKWNTLLIPPTTILINMGVWVAGTSYAINSGVQSGWPLFGKLPFAFWIVHLVYFF 1028
Db 845 ELYIYKWTLLIPPTTSLINMGVWVAGTSYAINSGVQSGWPLFGKLPFAFWIVHLVYFF 904
Query 1029 LKGLMGRQNRPTTIVVWAVLLASIFSLDLWRVDPFTTLAGPNIOQTGCINC 1080
Db 905 LKGLMGRQNRPTTIVILWSILLASVPSLVWVRINPFSVK---TDTTSLSLNC 953
RESULT 10
Ti0800
cellulose synthase (EC 2.4.1.-) catalytic chain celsA2 - upland cotton (fragment)
C:Species: Gossypium hirsutum (upland cotton)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10800
R:Pear, J.R.; Kawagoe, Y.; Schreckengost, W.E.; Delmer, D.P.; Stalker, D.M.
Proc. Natl. Acad. Sci. U.S.A. 93, 12637-12642, 1996
A:Title: Higher plants contain homologs of the bacterial celsA genes encoding the cataly
A:Reference number: Z17152; MUID:97057296; PMID:8901635
A:Accession: T10800
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-685 <PEA>
A:Cross-references: EMBL:U58284; NID:gi1706957; PIDN:AAB37767.1; PID:gi1706958
A:Experimental source: strain Acala SJ-2; fiber
C:Genetics:
A:Gene: celsA2
C:Function:
A:Description: involved in synthesis of cellulose
C:Keywords: glycosyltransferase; hexosyltransferase
Query Match 48.7%; Score 2816; DB 2; Length 685;
Best Local Similarity 74.0%; Pred. No. 2.6e-198;

Query Match	48.7%;	Score 2816;	DB 2;	Length 685;
Best Local Similarity	74.0%;	Pred. No. 2.6e-198;		

Db 712 EAJAVISCVEDNTEWGDRIWGYVTEVDVVTGYRMHNRGWSVYCTTKRDAFRGTAPI 771
QY 811 NLSRLNQLRNLWALGSEVIELPSRCHPLWYGGRLKFLERFAYINTIYPLTSLPLLYVC 870
Db 772 NLTDRLHQVLRWATGSVEIFFSKNAMF--ATRRLKFLQRLVAVLNVGIVPFTSIFLYVC 829
QY 871 ILPAICLLCTGKFMPEISNLASTWIALFLSIPAT---GILEMRWSGVGIDEWWRNEQF 926
Db 830 FLPALCLFSKGFVQSJ---DIHFSYLCITVTLTSLLELVKNSGIGLEEWWRNEQF 885
QY 927 WVIIGGSAHLFAVFGQLKLVLGADTNTFTVTSKANDEGD--FAELYMFKWTLTLLIPPT 984
Db 886 WLIGGTSALAAVVQGLLKVLAGIEISFTLTSKASGEDEDDIADLYIVKWTGLFIMPLT 945
QY 985 ILLINMVGWAGTSYAINSGYQSGPLFGKLFAPFWIVHLYPLKGLMGRQRTPTIVI 1044
Db 946 IIVNLVAIVIGASRTIYSVPOWGLMGIGIFFSLWLTHTMYPAKGLMGRGRKVTIYV 1005
QY 1045 VWAVLLASIFSLWVRVDP 1063
Db 1006 VNSGLVSIITVSLWITISP 1024

RESULT 15
C86446
probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
R;Theologits, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86446
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-979 <S70>
A;Cross-references: GB:AE005172; MID:gl0801364; PIDN:AAG23436.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

Query Match 37.4%; Score 2163.5; DB 2; Length 979;
Best Local Similarity 44.3%; Pred. No. 3.4e-150;
Matches 450; Conservative 173; Mismatches 245; Indels 147; Gaps 24;

QY 142 GEIPRGVPSVTSQMSGEIPGASPDHMHMSPTGNISRRAPFPYVNHSPNPSRFGSGIG 201
Db 4 GESPLRH-PRISHVNSGSDPGSSSDYNYK-----LVQ1PPTP--DNNPGPA-----SLS 50

QY 202 NVAMKERVDCWKQDKGAIPMTNGTSTAPSEGR-----AATDIDASTEYNMEDALLNDET 257
Db 51 IVL-----LEIDSNQESVPSVDIVSGSGKDNPEPLTDVIRNVGEEEDDTLLSKI 103

QY 258 RQPLSRKVPITASSKINPYRMVILVRLVLSIFLHYRLTNPVNRNAYPLWLLSVICELWAF 317
Db 104 SYSTRVVKISPIIALYRLIIVRVVSLALFLWRINPNKALWLLSVICELWAF 163

QY 318 SWILDQPKWFPINRETYDLRLALRY-----DREGEPSQLAAVDIFVSTVDPLKEPPIVT 372
Db 164 SWLLDQIPKLPFPVNHATDIEALKATFETPNPDNPTGKSDLPGLDVFVSTADAKEPEPLVT 223

QY 373 ANTVLSTILAVDYPVDKVSVDGASMLTDFALAESEFARKWVPFVKYDIEPRAPEF 432
Db 224 ANTILSILSVDPYVEKLSVYISDDGSLVTFEIAEAAEASFAKIWVPFCRKHIEPRNPES 283

QY 433 YFCQKIDYLDKQVQPSFVKDRRAMKREYEEFKIRNALV-----SK----- 473

Search completed: August 23, 2004, 01:09:14

Job time : 35 secs

Db 284 YFGKRDYKDVHRHDFVRRRYVKRAYDEFKVRVNALPHSIRRSRSDAFNSKEEIKALEK 343
QY 474 -----ALKVPEEGWIMODGTPWPG-----NNTROHPGMIQVFL- 506
Db 344 WKHKVKKVEBDQIKPRPALVAPKATW-MSDGTHTWPGTWAIVSGPHHSRGRDHASVIQVLLD 402
QY 507 -----GHSGGJLDTGNE--LPLRVVVSREKRPQFQHHKAGAMNALVRVSALVT 553
Db 403 PGDPEVEGKGGRALDLEGVDIRLPMLVIVSREKRPQFQHHKAGAMNALVRASAIMS 462
QY 554 NGQYMLNLDCDHYINNSKAVREAMCFMLDPNLGPQVCYQFQRPQFQDIDRNDRYANRNTV 613
Db 463 NGFFILNLDCDHVYNSRAFRDGCIFMMDHD-GDRVSYVQFQRPQFEGIDPSDRYANKTV 521
QY 614 FFDINRLDGIQGPVTVGTGCVNFNTAIYGYEPP-----IKAKPGFLASICGGKKKASKS 670
Db 522 FFDINRLDGIQGPVTVGTGCLFRRTALYGFNPPDVFVVEEFP--GSY-----FPLI 574
QY 671 KKRSSDKKSKNHVDSSVPVFNLEDEEGVEGAGFDDKSVLMSQMSLEKRFQSGSAFVA 730
Db 575 KKRSPATVAS-----EPEYTDDEDRFDIGL-IRKQFGSSSMLVN 613
QY 731 ST-LMEYGGVPOSSST-----PESL-----LKEAIIHVISCGYEDKSEWGTETI 770
Db 614 SVKVAEPGRPLATVHSSRLGRPPGSLTGSRKPLDPATVNEAVNVISCWYEDKTEWGFNV 673
QY 771 GWIYGSVTEDLITGFKMHARGWSVYCMKRPAPKGSAPINLSDLNQLRNLWALGSEVIEL 830
Db 674 GWIYGSVTEVDVVTGFRMHKGRWSFYCVTBPDAFRGSAPINLTDRLHQVLRWATGSVEIF 733
QY 831 FSRHCPWLWYGVGRLKFLERFAYINTIYPLTSLPLLYVCILPAICLLCTGKFMPEISNL 890
Db 734 FSRNNAIF--AGPKLKLQRLIQLNVGIVPFTSIFILTYCLPLPLSFGHFVVEITGS 791
QY 891 ASIWFIALLFLSIFATGILEMRWSGVGIDEWWRNEQFWVIGISAHLEFAVFGQLKLVLGAGI 950
Db 792 FLIYLLIITLSLCLGLAVLEVKWSGISLEEWWRNEQFWLIGTSAHLVAVLQGLKLVIAVG 851
QY 951 DTNFTVTSKA-----NDEGDPFAELYMFKWTLTLLIPPTIILINMVGWAGTSYAINSGYQ 1006
Db 852 EISFTLTSSKSTGGDEDEDFADLYLFKWTALMIPPLTIIILNIVAILFAVCRTVFSANP 911
QY 1007 SWGFLFKGLFFAFWVIVHLYPLKGLMGRQRTPTIIVWAVLLASIFSLWVRV 1061
Db 912 QWSNLLGGTFFASWVLLHMYPPFAKGLMGRGKPTVTVVWSGLIACLSLLIYITI 966

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2004, 22:25:29 ; Search time 19 Seconds
(without alignments)
2959.777 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MDGDADALKSGRHGAGDVQC.....VDPFTRLAGNIQTGGINC 1080

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	264	4.6	874	1	BCSA_SALTI
2	264	4.6	874	1	BCSA_SALTY
3	259.5	4.5	1596	1	ACG2_ACEXY
4	258	4.5	872	1	BCSA_ECO57
5	257	4.4	872	1	BCSA_ECOLI
6	248	4.3	739	1	BCSA_PSEFL
7	238.5	4.1	754	1	BCA1_ACEXY
8	238.5	4.1	1518	1	BCA2_ACEXY
9	238.5	4.1	1518	1	BCA5_ACEXY
10	237.5	4.1	756	1	BCA2_ACEXY
11	230	4.0	729	1	BCSA_XANAC
12	207	3.6	1550	1	ACG1_ACEXY
13	202.5	3.5	745	1	BCA3_ACEXY
14	119	2.1	1433	1	SUBF_BACSU
15	116.5	2.0	441	1	YDQJ_ECOLI
16	112.5	1.9	322	1	NUIM_STRPU
17	112	1.9	1029	1	RIP3_RAT
18	111	1.9	414	1	YOAB_BACSU
19	110.5	1.9	1154	1	KDGD_MESAU
20	110	1.9	753	1	STL5_RAT
21	109.5	1.9	1024	1	RIP3_MOUSE
22	108.5	1.9	1262	1	TPP2_MOUSE
23	108	1.9	1167	1	CLAA_BACSU
24	108	1.9	1894	1	BGS2_SCHPO
25	107.5	1.9	775	1	ICP0_HSV11
26	107.5	1.9	920	1	DMF1_SCHPO
27	107.5	1.9	1671	1	DPOL_PVRKO
28	107.5	1.9	3033	1	POLG_HCVJ6
29	107	1.9	730	1	STL5_HUMAN
30	107	1.9	1249	1	TPF2_RAT
31	106	1.8	357	1	G6PT_CANFA
32	106	1.8	1177	1	A11B_HUMAN
33	105.5	1.8	662	1	CYOB_BUCAI

34	105.5	1.8	886	1	YKR1_CABEL	P34307 caenorhabdi
35	105	1.8	357	1	G6PT_HUMAN	P35575 homo sapien
36	104.5	1.8	541	1	COX1_PODAN	P20681 podospora a
37	104.5	1.8	694	1	REP3A_HUMAN	Q9Y2J0 homo sapien
38	104.5	1.8	992	1	UVRA_MICLU	P13567 micrococcus
39	104	1.8	374	1	MTLD_BACHD	Q9K681 bacillus ha
40	104	1.8	495	1	AMPA_LEPIN	Q8F0Q1 leptospira
41	104	1.8	517	1	COX1_PARLI	P12700 paracentrot
42	104	1.8	1249	1	TPF2_HUMAN	P29144 homo sapien
43	103.5	1.8	530	1	CG2B_DROME	P20439 drosophila
44	103.5	1.8	622	1	COX1_BACSU	P24010 bacillus su
45	103.5	1.8	911	1	NIA_LYCES	P17570 lycopersico

ALIGNMENTS

RESULT 1
BCSA_SALTI STANDARD; PRT; 874 AA.
AC 082291, 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN BCSA OR STY4181 OR T3898.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RA MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
[2]
RP SEQUENCE FROM N.A.
RN STRAIN=Ty2 / ATCC 700931;
RC MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
uridine 5'-diphosphate glucose to cellulose, which is produced as
an extracellular component for mechanical and chemical protection
at the onset of the stationary phase, when the cells exhibit
multicellular behavior (rdar morphotype). Co-expression of
cellulose and thin aggregate fimbriae leads to a hydrophobic
network with tightly packed cells embedded in a highly inert
matrix (By similarity).
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
+ {(1,4)-beta-D-glucosyl}(N+1).
CC -!- COFACTOR: Magnesium (By similarity).
CC -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
(c-di-GMP) (By similarity).
CC -!- PATHWAY: Bacterial cellulose biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(Potential).
CC -!- DOMAIN: There are two conserved domains in the globular part of
the protein: the N-terminal domain (domain A) contains
the conserved DXD motif and is possibly involved in catalysis and

"Analysis of the Escherichia coli genome. V. DNA sequence of the RT region from 76.0 to 81.5 minutes.";
 RL Nucleic Acids Res. 22:2576-2586(1994).
 RN [2]
 RP REVISIONS.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP CHARACTERIZATION.
 RC STRAIN=ECOR10, ECOL12, and TOB1;
 RX MEDLINE=21160181; PubMed=11260463;
 RA Zogaj X., Nimtz M., Rohde M., Bokranz W., Roemling U.;
 RT "The multicellular morphotypes of Salmonella typhimurium and
 RT Escherichia coli produce cellulose as the second component of the
 RT extracellular matrix.";
 RL Mol. Microbiol. 39:1452-1463(2001).
 CC -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
 CC uridine 5'-diphosphate glucose to cellulose, which is produced as
 CC an extracellular component for mechanical and chemical protection
 CC at the onset of the stationary phase, when the cells exhibit
 CC multicellular behavior (rdar morphotype). Co-expression of
 CC cellulose and thin aggregative fimbriae leads to a hydrophobic
 CC network with tightly packed cells embedded in a highly inert
 CC matrix.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-beta-D-glucosyl)(N) = UDP
 CC + ((1,4)-beta-D-glucosyl)(N+1).
 CC -!- COFACTOR: Magnesium (By similarity).
 CC -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
 CC (C-di-GMP).
 CC -!- PATHWAY: Bacterial cellulose biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -!- DOMAIN: There are two conserved domains in the globular part of
 CC the protein: the N-terminal domain (domain A) contains the
 CC conserved DXD motif and is possibly involved in catalysis and
 CC substrate binding. The C-terminal domain (domain B) contains the
 CC QXXRW motif and is present only in processive glycosyl
 CC transferases. It could be involved in the processivity function of
 CC the enzyme, possibly required for holding the growing glycan chain
 CC in the active site.
 CC -!- MISCELLANEOUS: The genes bcsA, bcsB, bcsZ and bcsC are
 CC constitutively transcribed but cellulose synthesis occurs only
 CC when adrA, a putative transmembrane protein regulated by agfD, is
 CC expressed. Cellulose production is abolished in E.coli K12.
 CC -!- SIMILARITY: Belongs to the glycosyltransferase family 2.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 128.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U00039; AAB18510.1; ALT FRAME.
 CC EMBL: U00039; AAB18511.1; ALT FRAME.
 CC EMBL: AE000430; AAC76558.1; ALT_INIT.
 CC EcoGene: EGI2260; bcsA.
 CC InterPro: IPR003919; Cell_synth_A.
 CC InterPro: IPR001173; Glyco_trans_2.
 CC Pfam: PF00535; Glycos transf 2; 1.
 CC PRINTS: PR01439; CELL5NPHASEA.
 KW Cellulose biosynthesis; Transferase; Glycosyltransferase;
 KW Transmembrane; Inner membrane; Complete proteome.
 FT DOMAIN 271 364 CATALYTIC SUBDOMAIN A.

FT	DOMAIN	441	501	CATALYTIC SUBDOMAIN B.
FT	TRANSMEM	30	50	POTENTIAL.
FT	TRANSMEM	151	171	POTENTIAL.
FT	TRANSMEM	173	193	POTENTIAL.
FT	TRANSMEM	230	250	POTENTIAL.
FT	TRANSMEM	525	545	POTENTIAL.
FT	TRANSMEM	547	567	POTENTIAL.
FT	TRANSMEM	592	612	POTENTIAL.
FT	TRANSMEM	640	660	POTENTIAL.
FT	TRANSMEM	668	688	POTENTIAL.
FT	TRANSMEM	833	853	POTENTIAL.
FT	ACT SITE	313	313	POTENTIAL.
FT	ACT SITE	457	457	POTENTIAL.
FT	BINDING	360	360	SUBSTRATE (POTENTIAL).
FT	BINDING	362	362	SUBSTRATE (POTENTIAL).
SQ	SEQUENCE	872 AA;	99784 MW;	1432688A2B228F7 CRC64;

Query Match 4.4%; Score 257; DB 1; Length 872;
 Best Local Similarity 18.8%; Pred. No. 2e-10;
 Matches 156; Conservative 86; Mismatches 212; Indels 374; Gaps 29;

QY	263	RKVPTASSKINPYRMVIVLRVLVLSIFLHYRLT-----NPNVNAFPLMLLSVLCIWPFA	316
DB	191	RMPCGRFSAL----MLIVLSLTVSCRYTWRYTSTLNWDDPVSIVCGLLILFAETYAMIV	246
QY	317	LSWILDQPKWPPINRETVLDRLALRYDREGSPSOLAADVIEFVSTVDPLKEPPIVANTV	376
DB	247	L--VLGVQVWVFLNRQ-----VPLPKDMSLWPS---VDIFVPTYN---EDLNVVKNTI	293
QY	377	LSILAVDVPDKVSCYSDSDGASMLTFDALAETSEFARKWVFFVKKYDIEPRAPEFYFQC	436
DB	294	VASLGIDWPDKLNIWILDDG-----	315
QY	437	KIDYLDKQVQSFVKDRAMKREYEEFKIRINALYSKALKVPEEGWIMQDGTWPQGNTR	496
DB	316	-----REEFRQAFQNVG-----	327
QY	497	DHPGMIQVFLGHSGGLDTEGNELPRLVSVSREKRPGFQHHKAGAMNALVRVSAVLINQ	556
DB	328	-----VKIARTT-----HEHAKAGINNALKVA---KGE	354
QY	557	YMLNLDCDHYNNSKAVREAMCFMLDNLGPOVCVQFPQRF---DGIDRN---DRVAN	609
DB	355	FVSIQDCDHPVTRSPLOMTGMWFLKE---KQLAMQTPHHFFSPDPERNLGRFRKPTN	410
QY	610	RNTVFEDINLGLGICQGVVYVGTCCVNRRTAIYVEPPIKAKKPGFLASLGGKKASK	669
DB	411	EGTLFYGLVQDGNMDWDTFFCGSCAVIRR-----KP-----	442
QY	670	SKKRSSDKKKNKHVDSVPVFNLEDIEEGVEGAGFDDDEKSVLMSQMSLEKRFQGSAAV	729
DB	443	-----	442
QY	730	ASTLMEYGVQGSSTPESLLKEAIIHIVSGYEDKSEWGTETGMIYGVSVTEILITGFKMHA	789
DB	443	---LDEIGGI-----AVE-----TWTEAHTSLRLHR	466
QY	790	RGWRSVYCMKRPAPKGSAPINLSRLNQRLWALGSVELLSFRHCHPLWYGVGGRKFLF	849
DB	467	RGYTSAYM--RIPQAAGATSELSAHIGQRIRWARGMVQI--FRLDNPL--TGKGLKFAQ	520
QY	850	RFAYINTTYPLTSLPLIVYCIPLPAICLLTGKFMPEISNLASIFALFISIFATGILE	909
DB	521	RLCYVNAFMHFLSGIFRLIFLTAPLAFLLHAYI-----IYAPALMIALF-----	565
QY	910	MWRSVGIDEWNRNEQFWIGISAHLEFAVFGQLLKVLAGIDTNTVTTSKANDEGDFAE	969
DB	566	-----VLFPHM-----IHASL--TNSKIQGY--RHSFWSE	591
QY	970	LYMEKWTLLIIPPTTILLIINMGVVAGTSYAINSGYQSGWGLPFGKL-----PFA	1018
DB	592	IYETVLAWYIAPPTLVALIN-----PHKGFENTAKGLVEEYIV	631

QY 1019 FWIVHLVPLKGMGRQRTETIVW-----AVLLASIFSLMW 1059
 Db 632 DWISRPYIEVLVLL-----NLGVAVGLWIRYFGPTTEMLTVVWSWVW 675

RESULT 6

BCSA_PSEFL STANDARD; PRT; 739 AA.
 AC PS8931; Q8R8Z1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
 GN BCSA OR WSSB

OS Pseudomonas fluorescens.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_TaxID=294;

RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=SBW25;

RX MEDLINE=22013850; PubMed=12019221;

RA Spiers A.J., Kahn S.G., Bohannon J., Travisano M., Rainey P.B.;
 "Adaptive divergence in experimental populations of pseudomonas
 RT fluorescens. I. Genetic and phenotypic bases of wrinkly spreader
 RT fitness".
 RL Genetics 161:33-46(2002).

CC -1- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
 CC uridine 5'-diphosphate glucose to cellulose, which is produced as
 CC an extracellular component responsible for the structural
 CC integrity and rigidity of self-supporting mats characteristic of
 CC the "wrinkly spreader" phenotype.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-beta-D-glucosyl)(N) = UDP
 CC + ((1,4)-beta-D-glucosyl)(N+1).
 CC -1- COFACTOR: Magnesium (By similarity).
 CC -1- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
 CC (c-di-GMP) (By similarity).
 CC -1- PATHWAY: Bacterial cellulose biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (potential).

CC -1- DOMAIN: There are two conserved domains in the globular part of
 CC the protein: the N-terminal domain (domain A) contains
 CC the conserved DXD motif and is possibly involved in catalysis and
 CC substrate binding. The C-terminal domain (domain B) contains the
 CC QXRW motif and is present only in processive glycosyl
 CC transferases. It could be involved in the processivity function of
 CC the enzyme, possibly required for holding the growing glycan chain
 CC in the active site.

CC -1- SIMILARITY: Belongs to the glycosyltransferase family 2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AY074776; AAL71842.1; -

DR InterPro; IPR003919; Cell_synth A.

DR InterPro; IPR001173; Glyco_trans 2.

DR Pfam; PF00535; Glycos_transf_2; 1.

DR PRINTS; PR01439; CELLSYNHASEA.

KW Cellulose biosynthesis; Transferase; Glycosyltransferase;

Transmembrane; Inner membrane.

FT DOMAIN 157 250 CATALYTIC SUBDOMAIN A.

FT DOMAIN 327 387 CATALYTIC SUBDOMAIN B.

FT TRANSMEM 36 55 POTENTIAL.

FT TRANSMEM 59 76 POTENTIAL.

FT TRANSMEM 83 101 POTENTIAL.

FT TRANSMEM 116 138 POTENTIAL.

FT TRANSMEM 417 436 POTENTIAL.

FT TRANSMEM 440 462 POTENTIAL.

FT TRANSMEM 524 546 POTENTIAL.
 FT TRANSMEM 551 573 POTENTIAL.
 FT ACT_SITE 199 199 POTENTIAL.
 FT ACT_SITE 343 343 POTENTIAL.
 FT BINDING 246 246 SUBSTRATE (POTENTIAL).
 FT BINDING 248 248 SUBSTRATE (POTENTIAL).
 SQ SEQUENCE 739 AA; 82165 MW; 2B962EA3854B23BB CRC64;

Query Match 4.3%; Score 248; DB 1; Length 739;

Best Local Similarity 19.4%; Pred. No. 7.1e-10;

Matches 155; Conservative 85; Mismatches 212; Indels 348; Gaps 34;

QY 263 RKVPIASSKINPYRMVIVLRLVLSI-----FLHYRLTNPNVRYAPLWL-----LSVICE 312

Db 77 RKIP-----GRLLALVLLVLSVASLRYFWRLTSTL--GFETWDMFQYGLVAA 125

QY 313 INFAL-----SWLDQFPKWFPEINRETYLDRLALRYDREGESQAAVDIFVSTV 362

Db 126 EFVALIVLIEGYVQTAWPLRRTPWL-----KTEPEEWPTVDVFITY 168

QY 363 DPLKEPIVTANTVLSILAVDYPVDKVCYVDDGASMLTFDALAETSEFARKWVPFVK 422

Db 169 N---EALSIVKLTIFAQAQMDPKDKLRVHVLDDG----- 200

QY 423 YDIEPRAPEF-YFCOKIDYLDKVPQSFVKDRAMKREYEFEFKIRINALYSKALKVPEEG 481

Db 201 -----RDDDFREFCRKV----- 212

QY 482 WIMQDGTPEFNNTRDHPGMIQVPLGHSGGLDTEGNELPLRVVYSREKRGFQHHKAGA 541

Db 213 -----GVN-----YIRDN-----NEHAKAGN 229

QY 542 MNALVRYSVLTNGOYMLNLDCHYINNSKAVREAMCFMDNPLGPOVCVYQVQRF 598

Db 230 LINEALKY-----TDGEYIALEFDADHVPTRSFQLVSLGFWFLKDKPL-----AMLQTPHFFSP 281

QY 599 DGDIDRN-DRY---ANRNTVFDDNLRLGDLGIQGVVYGTGCVFNRTAIYGYEPPKAKP 654

Db 282 DPEKNUDTFRVNPNEGELFYGLVQDNDLWNATFFCGSCAVIR-----EP----- 328

QY 655 GFLASLGGKKKASKSKKSSDKKSNKHVDSSVPVFENLEDIEBEGVAGDFDDEKSVLMS 714

Db 329 ----- 328

QY 715 QMSLEKRFQSGAAFAVASTLMYEGVQVQSPSTPESLLKEAIIHVISCYEDKSEWGTEIGWIY 774

Db 329 -----LLEIGGV-----AVE----- 338

QY 775 GSVTDELITGFMHARGMRSVYCMKRPAPFKGSAPINLSDRLNOLRWALGSVILFSRH 834

Db 339 -TVTEDAHTALKUNRLGINTAYL--AIPQAAGLATESLSRHINORIRWARGMAQI-PRTD 394

QY 835 CPLWYGGYGRKLFELERFAYINTTIYPLTSLPLVYCYLPAICLLTGKFIEMPEISLASIW 894

Db 395 NPL---LGKGLKWQQRICYANAMQHFFYGLPRLVFLTAFLAYLIFG---AEIFHASALM 447

QY 895 FIALFL-SIFATGILENRWSGVGIDENWRNEQF-----WVIGGISAHLEFAVQGLLKVLA 948

Db 448 IVAYVPLPHLVHSSLTNSRIQGRFRHSFW-NEVYETVLAWYI-----LPPVLVALVNPKA 500

QY 949 GIDTNFTVTSKANDEEDGFARLYMFKWTTLLIPPTTILINMVGWVAGTSYAINSGYQSW 1008

Db 501 G---GFNVTDK-----GGIDKQFFDW-KLARPYVLLAVNLIGLGFGIHQLI-----W 545

QY 1009 GPLFGKLFAPF---WVIVHL 1025

Db 546 GDASTAVTVAINLTWLTLYNL 565

RESULT 7

BCAL_ACEXY

ID_BCAL_ACEXY STANDARD; PRT; 754 AA.

AC P19449;

DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
 GN BCSA.
 OS Acetobacter xylinus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
 OC Acetobacteraceae; Gluconacetobacter.
 OX NCBI_TaxID=28448;
 RN [1]__
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-42 AND 196-206.
 RC STRAIN=1306-3;
 RX MEDLINE=91045951; PubMed=2146681;
 RA Wong H.C., Fear A.L., Calhoun R.D., Eichinger G.H., Mayer R.,
 RA Amikam D., Ben-Ziman M., Gelfand D.H., Meade J.H., Emerick A.W.,
 RA Bruner R., Ben-Bassat A., Tal R.;
 RT "Genetic organization of the cellulose synthase operon in Acetobacter
 RT xylinum";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8130-8134 (1990).
 RN [2]
 RP ENZYME REGULATION.
 RC STRAIN=1306-3;
 RX MEDLINE=21194671; PubMed=11297407;
 RA Chang A.L., Tuckerman J.R., Gonzalez G., Mayer R., Weinhouse H.,
 RA Volman G., Amikam D., Ben-Ziman M., Gilles-Gonzalez M.-A.;
 RT "Phosphodiesterase A1, a regulator of cellulose synthesis in
 RT Acetobacter xylinum, is a heme-based sensor";
 RL Biochemistry 40:3420-3426 (2001).
 CC -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes
 CC uridine 5'-diphosphate glucose to cellulose in a processive way.
 CC The thick cellulosic mats generated by this enzyme probably
 CC provide a specialized protective environment to the bacterium.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-beta-D-glucosyl}(N) = UDP
 CC + {(1,4)-beta-D-glucosyl}(N+1).
 CC -!- COFACTOR: Magnesium.
 CC -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid
 CC (c-di-GMP).
 CC -!- PATHWAY: Bacterial cellulose biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -!- INDUCTION: Cellulose is produced at a linear rate with respect to
 CC cell growth when O(2) is present.
 CC -!- DOMAIN: There are two conserved domains in the globular part of
 CC the catalytic subunit: the N-terminal domain (domain A) contains
 CC the conserved DXD motif and is possibly involved in catalysis and
 CC substrate binding. The C-terminal domain (domain B) contains the
 CC OXXRW motif and is present only in processive glycosyl
 CC transferases. It could be involved in the processivity function of
 CC the enzyme, possibly required for holding the growing glycan chain
 CC in the active site.
 CC -!- SIMILARITY: Belongs to the glycosyltransferase family 2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M37202; AA21884.1; -.
 DR PIR; A43735; A43735.
 DR InterPro; IPR003919; Cell synth A.
 DR InterPro; IPR001173; Glyco trans 2.
 DR Pfam; PF00535; Glycos transf 2; 1.
 DR PRINTS; PR01439; CELLSYNTHASEA.
 DR Cellulose biosynthesis; Transferase; Glycosyltransferase;
 KW Transmembrane; Inner membrane.
 FT DOMAIN 147 240
 FT DOMAIN 317 377 CATALYTIC SUBDOMAIN A.
 FT TRANSMEM 26 46 CATALYTIC SUBDOMAIN B.
 FT TRANSMEM 47 67 POTENTIAL.
 FT TRANSMEM 108 128 POTENTIAL.

FT	TRANSMEM	167	187	POTENTIAL.
FT	TRANSMEM	407	427	POTENTIAL.
FT	TRANSMEM	430	450	POTENTIAL.
FT	TRANSMEM	468	488	POTENTIAL.
FT	TRANSMEM	516	536	POTENTIAL.
FT	TRANSMEM	549	569	POTENTIAL.
FT	ACT SITE	189	189	POTENTIAL.
FT	ACT SITE	333	333	POTENTIAL.
FT	BINDING	236	236	SUBSTRATE (POTENTIAL).
FT	BINDING	238	238	SUBSTRATE (POTENTIAL).
SQ	SEQUENCE	754 AA;	84442 MW;	8DSFCIF62E2C068 CRC64;

Query Match 4.1%; Score 238.5; DB 1; Length 754;
 Best Local Similarity 20.1%; Pred. No. 3.6e-09;
 Matches 163; Conservative 82; Mismatches 232; Indels 335; Gaps 32;

QY	279	IVLRVLVLSIFLHYRLTNPVRNAYPLWLLSVICETWFLSWILDQ---PPKWFP-----	329
DB	53	LIVLVLCVLFVFIYGRGKSRRTQIFLEVLISALVSLRY-LTWRLTETLDFDTWIOGGLGVT	111
QY	330	-INRETY-LDLRLALRYDREGPSQLA-----AVDIFVSTVDPLKEPPIVTANT	375
DB	112	LIMAEYALYMLFLSYFQTIQLHRAPLPLPDNDVDWPTVDIFITYD---EQLSIVRLT	168
QY	376	VLSILAVDYPVDKVCYVSDGASMLTFDALAETSEFARKWVPFVKKYDIEPRAPEFYFC	435
DB	169	VLGALGIDWPPDKVNVYILDG-----	190
QY	436	QKIDYLDKQVPSFVKDRRAMKREVEEPKIRINALVSKALKVPBERGIMQDTPPGNNT	495
DB	191	-----VRP-----EPEQFAKDCGAL-----	205
QY	496	RDPGMIQVFLGHSGGLDEGNEPLRLVYVSRKPGFHGHKAGAMNALVRVSAVLING	555
DB	206	-----YIGRVDSS-----HAKAGNLNHAIK-----RUSG	229
QY	556	QYMLMLDCDHYNNSKAVREACFLMDNLPQVCYQVPPQRFQDFGIDRNDRYANENTVFF	615
DB	230	DYILLDCDH-IPTRAFQLIANGWVADR---KIALMQTPHHFYSPD-----	272
QY	616	DINLRGLDGIQGPVYVGTGVFNRTAIYGYBPPIKAKPGFLASLCGGKKKASKSKRS	675
DB	273	-----PFQRLAVGYRTP-----	285
QY	676	DKKSNKHVDSSVPVFNLEDEEGVEGAGFDEKSVLMSQMSLEKRFQSAAFVASTIME	735
DB	286	-----PEGNL---FYGVIQDND-----FWDATFFCGSCAI- 313	
QY	736	YGGVPQSSTPESLKEAIIHVISCGYEDKSEWGTETIGWYGVSTEDILTGFKNHARGWSV	795
DB	314	-----LRRAIESIG-----GPAVETVTEDAHTALRMQRGWSTA 348	
QY	796	YCMPKRPAPKSAFINLDRNLQVLRWALGSVEILFSRHCLWYGYGRLLKPLERFAYIN	855
DB	349	YL--RIPVASGLATERLTHIGQRMWARGMIQI-PRVDNPM---LGGLGLQGLRLCYLS	402
QY	856	TTIYPLTSLPLVLCILPAICLLTGKFTMPBSNLASLWIALFLSIFATGILEMWS--	913
DB	403	AMTSEFFFAIPRIVFLASPLAFLPFGQNI-----AASPLAVLAYAIPMHFHSIA	451
QY	914	-GVGIDEWNRNEQFW-VIGGISAHLFAVFQGLLKLVLAGIDNTFTVTSKA---NDEGDFA	968
DB	452	TAQKVNKGWR-YSEWSEVYETTMALFLVRVTIITLPPSKGKFNVEKGVLEEEFDLG	510
QY	969	ELYMEKWTLLIPTTILLINWGVVAGTSVAINSGYQSWGPLFG--KLFFAFWIVVHLY	1026
DB	511	ATY-----PNIIF-----AGIMTLGLLGLFLFTFHF-----	537
QY	1027	PFLKGLMGRQNRTPTIWVA-----VLLAST 1053	
DB	538	---NQLAGIAKRAVILNLCINWAMISILILLAAI 566	

RP SEQUENCE FROM N.A., AND SEQUENCE OF 768-781.
RC STRAIN=ATCC 53582;
RA MEDLINE=91322509; PubMed=1830823;
RX Saxena I.M., Lin F.C., Brown R.M. Jr.;
RT "Identification of a new gene in an operon for cellulose biosynthesis
in Acetobacter xylinum.";
RL Plant Mol. Biol. 16:947-954 (1991).
RN [3]
RP REVISIONS.
RC STRAIN=ATCC 53582;
RX MEDLINE=94364954; PubMed=8083166;
RA Saxena I.M., Kudlicka K., Okuda K., Brown R.M. Jr.;
RT "Characterization of genes in the cellulose-synthesizing operon (acs
operon) of Acetobacter xylinum: implications for cellulose
RT crystallization.";
RL J. Bacteriol. 176:5735-5752 (1994).
RN [4]
RP SEQUENCE OF 1-8 FROM N.A.
RC STRAIN=ATCC 23769;
RX MEDLINE=94131945; PubMed=8300521;
RA Standal R., Iversen T.-G., Coucheron D.H., Fjaervik E., Blatny J.M.,
RT Valla S.;
RT "A new gene required for cellulose production and a gene encoding
cellulolytic activity in Acetobacter xylinum are colocalized with the
RT bcs operon.";
RL J. Bacteriol. 176:665-672 (1994).
RN [5]
RP FUNCTION.
RC STRAIN=ATCC 53582;
RX MEDLINE=90202941; PubMed=2138620;
RA Lin F.C., Brown R.M. Jr., Drake R.R. Jr., Haley B.E.;
RT "Identification of the uridine 5'-diphosphoglucose (UDP-Glc) binding
RT subunit of cellulose synthase in Acetobacter xylinum using the
RT photoaffinity probe 5-azido-UDP-Glc.";
RL J. Biol. Chem. 265:4782-4784 (1990).
RN [6]
RP 3D-STRUCTURE MODELING, AND MUTAGENESIS OF ASP-188; ASP-189; ASP-236;
ASP-333; GLN-369; ARG-370 AND ARG-372.
RC STRAIN=ATCC 23769;
RX MEDLINE=21324818; PubMed=11430986;
RA Saxena I.M., Brown R.M. Jr., Dandekar T.;
RT "Structure-function characterization of cellulose synthase:
RT relationship to other glycosyltransferases.";
RL Phytochemistry 57:1135-1148 (2001).
RN [7]
RP REVIEW ON DOMAIN ARCHITECTURE.
RC MEDLINE=95189716; PubMed=7883697;
RA Saxena I.M., Brown R.M. Jr., Fevre M., Geremia R.A., Henrissat B.;
RT "Multidomain architecture of beta-glycosyl transferases: implications
RT for mechanism of action.";
RL J. Bacteriol. 177:1419-1424 (1995).
CC -!- FUNCTION: Bifunctional protein comprised of a catalytic subunit
CC and a regulatory subunit. The catalytic subunit of cellulose
CC synthase polymerizes uridine 5'-diphosphate glucose to cellulose
CC in a processive way. The thick cellulosic mats generated by this
CC enzyme probably provide a specialized protective environment to
CC the bacterium. The regulatory subunit binds bis-(3'-5') cyclic
CC diguanylic acid (c-di-GMP).
CC -!- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-beta-D-glucosyl)(N) = UDP
CC + ((1,4)-beta-D-glucosyl)(N+1).
CC -!- COFACTOR: Magnesium (By similarity).
CC -!- ENZYME REGULATION: Activated by c-di-GMP.
CC -!- PATHWAY: Bacterial cellulose biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (potential).
CC -!- INDUCTION: Cellulose is produced at a linear rate with respect to
CC cell growth when O(2) is present.
CC -!- DOMAIN: There are two conserved domains in the globular part of
CC the catalytic subunit: the N-terminal domain (domain A) contains
CC the conserved DXD motif and is possibly involved in catalysis and
CC substrate binding. The C-terminal domain (domain B) contains the
CC OXXRW motif and is present only in processive glycosyl
CC transferases. It could be involved in the processivity function of

CC the enzyme, possibly required for holding the growing glycan chain
CC in the active site.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC GLYCOSYLTRANSFERASE FAMILY 2.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSE/BCSB
CC FAMILY.
CC -!- CAUTION: Was originally (Ref.1) thought to be two separate ORFs
CC named acsA and acsB, due to a frameshift in position 678.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X54676; CAA38487.1; ALT_FRAME.
CC EMBL; X54676; CAA38488.1; ALT_FRAME.
CC EMBL; M96060; AAA16971.1; -.
CC InterPro; IPR003919; Cell_synth_A.
CC InterPro; IPR003920; Cell_synth_B.
CC InterPro; IPR001173; Glyco_trans_2.
CC Pfam; PF031170; BCSB; 1.
CC Pfam; PF00535; Glycos_transf_2; 1.
CC PRINTS; PR01439; CELLSNTHASEA.
CC PRINTS; PR01440; CELLSNTHASEB.
CC Cellulose biosynthesis; Transferase; Glycosyltransferase;
CC Transmembrane; Inner membrane.
CC DOMAIN 1 741
CC CYCLIC DI-GMP BINDING.
CC DOMAIN 742 1550
CC CATALYTIC SUBDOMAIN A.
CC DOMAIN 147 240
CC CATALYTIC SUBDOMAIN B.
CC DOMAIN 317 377
CC POTENTIAL.
CC TRANSMEM 26 46
CC TRANSMEM 47 67
CC TRANSMEM 106 126
CC TRANSMEM 398 418
CC TRANSMEM 423 443
CC TRANSMEM 468 488
CC TRANSMEM 507 527
CC TRANSMEM 547 567
CC TRANSMEM 1513 1533
CC ACT_SITE 189 189
CC ACT_SITE 333 333
CC BINDING 236 236
CC BINDING 238 238
CC MUTAGEN 188 188
CC MUTAGEN 188 188
CC MUTAGEN 189 189
CC MUTAGEN 236 236
CC MUTAGEN 333 333
CC MUTAGEN 369 369
CC MUTAGEN 370 370
CC MUTAGEN 370 370
CC MUTAGEN 372 372
CC SEQUENCE 1550 AA; 168161 MW; 63AB8952BC39E961 CRC64;
CC
CC Query Match 3.6%; Score 207; DB 1; Length 1550;
CC Best Local Similarity 19.6%; Pred. No. 2.2e-06;
CC Matches 154; Conservative 88; Mismatches 245; Indels 298; Gaps 29;
CC
CC Qy 280 VLRVLSIFLHYELTNPRNAYPLWL-----LSVCEIWFALSWILDQPKWFPINRE 333
CC Db 80 VLSGVSIRLVTKLTETL--SFDTWQGLLGTMLLVLYALMMFLSVFTQTIAFLHRA 137
CC Qy 334 TYDLRLALRYDREGEPQSLAAVDIFSVTDPLKEPPIVTANTVLSILAVDPVDKVCYV 393
CC Db 138 P-----LPL-----PENDEWPTVDIFVPTYN--EELSIVRLTVLSGLGIDWPPEKVRVHI 186
CC Qy 394 SDDGASMLTFDALETSEFAKRVVFKKVDIEPRAPREFYFCQKIDYLDKVKQVPSFKVDR 453
CC Db 187 LDDG-----RRPEF----- 195

Db 247 QISGMVMSDS---NIALQTPHFYSPD----- 272
QY 634 GCVENRTAIVGYEPIKAKPGFLASLGGKKASKSKKSKKNDVSSVVPVNL 693
Db 273 --PFQRNLAVGYRTP-----PBGNL 290
QY 694 EDIEBEGVAGFDDKSVLMSQMSLEKRFSGSAFVASTLMEYGGVQSSPTPESLKKEAI 753
Db 291 ---FYGVIQDND-----FWDATFCGSCAI-----LRRKAI 319
QY 754 HVISCGYEDKSEWGTGVIWYGSYTEDILTGFKMHARGWSVYCMRPKPAFKGSAFINLS 813
Db 320 EEIG-----GPAETVTEDAHTALRMQRKWSYAYL--RIPLASGLATERLI 364
QY 814 DRLNQLVRWALGSVEILFSRHCPWYGVGRKLKELERPAYINTIYPLTSLPLLYCILP 873
Db 365 THIGQRMRWARGMTQI-FRVDNPM---LGSGLKLGQRCLYSAMTSFFFAIPRVIFLASP 420
QY 874 AICLLTGKFMPEISNLASTWFIALFLSIPATGILEMRWSGVGIDEMRNEQFW--VIGG 931
Db 421 LALFFSQNIIASPLAVGVAIPHMFHSIAT-----AAKNVKGWR-YSFVSEVYET 471
QY 932 ISAHFAVFOGLLKLAVIGADITNFTVTSKA---NDEEGDFAELYMFKWTTLLIPPTTILII 988
Db 472 VMA-LFLVRVTIVTMLPFSKGVNTEKGGVLEREEFDLTATY-----PNIIFALI 521
QY 989 NMGVVWAG-----TSYAINSGVQSGWPLFGKLFPAFWIVVHLYPFLKGL-M 1033
Db 522 MALGLRGLYALIFQHLDIISERAYALN-----CIWSVISLIILMAVISV 566
QY 1034 GRQNR 1038
Db 567 GRETK 571

RESULT 14

SUBF_BACSU
ID SUBF_BACSU STANDARD; PRT: 1433 AA.
AC P16397;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Bacillopeptidase F precursor (EC 3.4.21.-) (Esterase) (RP-I protease)
DE (90 kDa serine proteinase).
GN BPR OR BPR OR BSU15300.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 195-222.
RX MEDLINE=90170864; PubMed=2104512;
RA Sloma A., Rufo G.A. Jr., Rudolph C.F., Sullivan B.J., Theriault K.A.,
RA Pero J.;
RT "Bacillopeptidase F of Bacillus subtilis: purification of the protein
RT and cloning of the gene.";
RL J. Bacteriol. 172:1470-1477 (1990).
RN [2]
RP REVISIONS.
RC STRAIN=168;
RX MEDLINE=90368623; PubMed=2118514;
RA Sloma A., Rufo G.A. Jr., Rudolph C.F., Sullivan B.J., Theriault K.A.,
RA Pero J.;
RL J. Bacteriol. 172:5520-5521 (1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90216713; PubMed=2108961;
RA Wu X.-C., Nathoo S., Pang A.S.-H., Carne T., Wang S.-L.;
RT "Cloning, genetic organization, and characterization of a structural
RT gene encoding bacillopeptidase F from Bacillus subtilis.";
RL J. Biol. Chem. 265:6845-6850 (1990).
RN [4]
RP SEQUENCE FROM N.A.

RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Arevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holappert S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Toato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzneger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256 (1997).
RN [5]
RP SEQUENCE OF 1-211 FROM N.A.
RX MEDLINE=89008108; PubMed=3139638;
RA Beall B., Lowe M., Lutkenhaus J.;
RT "Cloning and characterization of Bacillus subtilis homologs of
RT Escherichia coli cell division genes ftsZ and ftsA.";
RL J. Bacteriol. 170:4855-4864 (1988).
RN [6]
RP SEQUENCE OF 1410-1433 FROM N.A.
RX STRAIN=168 / Marburg;
RC MEDLINE=90174995; PubMed=2106671;
RA Masuda E.S., Anaguchi H., Sato T., Takeuchi M., Kobayashi Y.;
RT "Nucleotide sequence of the sporulation gene spoIIIGA from Bacillus
RT subtilis.";
RL Nucleic Acids Res. 18:657-657 (1990).
RN [7]
RP SEQUENCE OF 195-219.
RC STRAIN=Natto 16;
RA Kato T., Yamagata Y., Arai T., Ichishima E.;
RT "Purification of a new extracellular 90-kDa serine proteinase with
RT isoelectric point of 3.9 from Bacillus subtilis (natto) and
RT elucidation of its distinct mode of action.";
RL Biosci. Biotechnol. Biochem. 56:1166-1168 (1992).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; M29035; AAA62679.1; -;
DR EMBL; J05400; AAA83362.1; -;
DR EMBL; Z99111; CAB13403.1; -;
DR EMBL; M22630; AAA22458.1; -;
DR EMBL; X17344; CAA35224.1; -;
DR PIR; A36734; A36734.

```
DR HSSP; P00782; 2SBT.
DR MEROPS; S08.017; -.
DR SubtilList; BG10233; bpr.
DR InterPro; IPR008757; Peptidase_M6.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR009020; Protease_inhib.
DR Pfam; PF05547; Peptidase_M6; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; FALSE_NEG.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR KW Hydrolyase; Serine protease; Zymogen; Signal; Complete proteome.
FT SIGNAL 1 30
FT PROPEP 31 194
FT CHAIN 195 755
FT PROPEP 756 1433
FT ACT_SITE 227 237
FT ACT_SITE 274 274
FT ACT_SITE 452 452
FT CONFLICT 219 219
FT CONFLICT 393 393
FT CONFLICT 829 834
FT CONFLICT 836 841
FT CONFLICT 844 852
FT CONFLICT 853 1433
SQ SEQUENCE 1433 AA; 154577 MW; 98DF6846897807C9 CRC64;

Query Match 2.1%; Score 119; DB 1; Length 1433;
Best Local Similarity 19.8%; Pred. No. 4.7; Indels 334; Gaps 49;
Matches 184; Conservative 96; Mismatches 314;

QY 13 HG---AGDVQCICADGLGTLTLDGVFTACDVCVFCPCVPCYBHER-----54
DB 503 HCLVNAFAVSAVTDGLGKA-EGQVSVEGDDEPPV---YQHKVTEAYEGSGSLPLTIL 557
QY 55 KEGTQACIQCKTKYKRRHG-----SPAIRBEGDDTDADGSDGFNYPASGTEDQKQKTA 108
DB 558 AEDNVSVTSVKLSYKLDQGEWTEITAKRISGDHLKGT-----YQAEIPDIKGTKLS 608
QY 109 DMRSWRMNTGSGNVGHKPYD-----SGEIGLSKYDSGEIPRGVVPVSTNSQMSGEIPG 163
DB 609 YK---WMTHDFGHVSSVDYDVTVKPSTAGY-KODFETAPGGWVAGSTNNWVGVPVS 664
QY 164 ASPDH-----HMSPTGNISRRAPPVYNHSPNPSREFSGSI-----200
DB 665 TCPNTAASGEKVGYNLTGNVANSANMLV--MPPIKAPDSGLFLQPKSWHNLEDDPDY 722
QY 201 -----GNVANK-----ERVD--GWKMKQDKGAIPMTNGTSTIAPSEGRA 236
DB 723 GYVFLPEGEKNWEQAGVYNGKTSSTWDEIDLSAYKQNIQVMFNLSQSDSIA-KEGWY 781
QY 237 ATDI---DASTEYNMEDALLNDETQPLSRKVPYIASSKINPYRMVILVLSIFLHYR 293
DB 782 IDVVLSKSAKTVKKNKLGVEKPSGQKKKPVNPKAKPSANTAVKH-----QNK 833
QY 294 LTNPVNRNAYPLWLLSVICIEIPALSWILDDQPKFPINRETLDRLALRY-----DREGPS 350
DB 834 ALQP--QVPLKAQVSVE-----TGKSTYSDOSTGGYTLKHAGDYT 874
QY 351 QLAADVIFVSTVD--PLKEPPIVTVANTVLSIL-----AVDVPVVKVCVSDSD 396
DB 875 LMAEAYGQSKTKVSLKTDOTQANFTLEEMKKGTLKGTINKTTGEPVTCASVYVVED 934
QY 397 GA--SMLTFDA--LAETSEFARKWVFFVKYVDIEPRAPEFYFCQKIDLYLKDQVPSFK 451
DB 935 AAVEPAMNDKGEYMLEAVEGA-----YTIKVAAPGYI-----967
QY 452 DRAMKREYEFKIRINALVSK--ALK--VPEEGWIMQD-----GTPWPGNN 494
DB 968 -----SDEFSVELKGDVTKETALKPVPVGPGEIAYDDGAENANSYFAAGNGWAVKM 1019
QY 495 T-----RDHPGMLQVFLHSGGL-----DTE-----GNEPLRLVY--VSREKRFQFHKKA 539
```


Db 61 CPQCKTKYKRHKGSFAIRGEEGEDTDADDVSDYNYPASGSADQKQKIADRMRSWRMAGG 120
Qy 121 SGNVCHPKYDSGEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMHMSPTGNISRR 180
Db 121 GGDVGRPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGEIPGASPDHMHMSPTGNIGKR 180
Qy 181 APFPVYNHSPNPSRFFSGISGNVANKERVYDGMKQDKGAIPTWNGTSTAPSEGRAATDI 240
Db 181 APFPVYNHSPNPSRFFSGISGNVANKERVYDGMKQDKGAIPTWNGTSTAPSEGRVGD 240
Qy 241 DASTYNNMEDALLNDETROPQSRKVPPIASSKINPYRMVIVLRLVLSIFLHRLNPNVRN 300
Db 241 DASTYNNMEDALLNDETROPQSRKVPPIASSKINPYRMVIVLRLVLSIFLHRLNPNVRN 300
Qy 301 AYPLMLLSVICIWFALSILAVDYPVKVSCVSDDGASMLTFDALAETSEFARKWVPV 360
Db 301 AYPLMLLSVICIWFALSILAVDYPVKVSCVSDDGASMLTFDALAETSEFARKWVPV 360
Qy 361 TVDPLKEPPIVTANTVLSILAVDYPVKVSCVSDDGASMLTFDALAETSEFARKWVPV 420
Db 361 TVDPMKEPPIVTANTVLSILAVDYPVKVSCVSDDGASMLTFDALAETSEFARKWVPV 420
Qy 421 KKYDTEPRAPYFQKIDYLDKQVQPSVVDKRRAMKREYEFKIRINALVSKALKVPBE 480
Db 421 KKYNTIEPRAPYFQKIDYLDKQVQPSVVDKRRAMKREYEFKIRINALVSKALKVPBE 480
Qy 481 GWIMODGTPWPGNTRDHPGMIQVFLHSGGLDTEGNELPRLVYVSREKRPQFQHHKAG 540
Db 481 GWIMODGTPWPGNTRDHPGMIQVFLHSGGLDTEGNELPRLVYVSREKRPQFQHHKAG 540
Qy 541 ANNALVRVSAVLNNGOYMLNDCDHYINNSKAVREAMCFMDPNLGPQVCYVQFPORFDG 600
Db 541 ANNALVRVSAVLNNGOYMLNDCDHYINNSKAVREAMCFMDPNLGRSVCIYVQFPORFDG 600
Qy 601 IDRDYANRNTVFFDINRLGLDGIQGPVYVGTGCVFNRTAIYGEPPITKAKPG-FLAS 659
Db 601 IDRDYANRNTVFFDINRLGLDGIQGPVYVGTGCVFNRTAIYGEPPITKQKKGSFLSS 660
Qy 660 LCGGKKKASKRSDDKKSNKNDSSVPVFNLEIDIEGVGAGFDDKSLVMSQMSLE 719
Db 661 LCGGKKKASKRSDDKKSNKNDSSVPVFNLEIDIEGVGAGFDDKSLVMSQMSLE 720
Qy 720 KRFQGSAAFAVSTLMEYGGVPOSSTPESLLKBAIHVISCYEDKSEWGTGEMWIGYSVTE 779
Db 721 KRFQGSAAFAVSTLMEYGGVPOSATPESLLKBAIHVISCYEDKTEWGTGEMWIGYSVTE 780
Qy 780 DILTGFKMEHARGMRSYVCMKRPAPKGSAPINLSRLNQVLRLWALGSVEILFSRHCPWY 839
Db 781 DILTGFKMEHARGMRSYVCMKRPAPKGSAPINLSRLNQVLRLWALGSVEILFSRHCPWY 840
Qy 840 GYGGRLKFLERAYINTTIYPLTSLPLLYCYILPATCLLTGKFIMPEISNLASIMFIALF 899
Db 841 GYGGRLKFLERAYINTTIYPLTSLPLLYCYILPATCLLTGKFIPPEISNFAISIMFISLF 900
Qy 900 LSFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHLFAVFQGLLKLVLGIDTNTFTVTSK 959
Db 901 ISIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHLFAVFQGLLKLVLGIDTNTFTVTSK 960
Qy 960 ANDEGDFAELMYFKWTTLLIPPTTLLIINMGVAVAGTSYAINSGVQSGPLFGKLPFAF 1019
Db 961 ASDEGDFAELMYFKWTTLLIPPTTLLIINMGVAVAGTSYAINSGVQSGPLFGKLPFAF 1020
Qy 1020 WVIVHLYPPLKGLMGQRNTPPTIVVAVILLASIFSLLMVRVDPFTTRLAGNIOQCGLN 1079
Db 1021 WVIVHLYPPLKGLMGQRNTPPTIVVAVILLASIFSLLMVRVDPFTTRVGTQTCGLN 1080
Qy 1080 C 1080
Db 1081 C 1081

RESULT 2
Q9LLI1

ID Q9LLI1 PRELIMINARY; PRT; 1079 AA.
AC Q9LLI1;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Cellulose synthase-9.
GN CESA-9.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
RN NCBI_TaxID=4577;
RX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xoonostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (CesA) gene
RT family.";
RL Plant Physiol. 123:1313-1324(2000).
DR EMBL; AF200533; AAF89969.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
DR GO; GO:0030244; P:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose_synth.
DR InterPro; IPR001841; Znf ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1; 10762 MW; EB334272C2ED2DD7 CRC64;
SQ SEQUENCE 1079 AA; 120162 MW; EB334272C2ED2DD7 CRC64;
Query Match 93.9%; Score 5423.5; DB 10; Length 1079;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 996; Conservative 46; Mismatches 37; Indels 1; Gaps 1;
Qy 1 MDCDADALSKRHGAGDVCOICADGLGTLTLDGVDVTACDVCPFCPCYEHERKEGTOA 60
Db 1 MEGDADGVKSGRRGGQVCQICGQGVGTGTTAGDVTACDVCPFCPCYEHERKDGTOA 60
Qy 61 CLQCKTKYKRHKGSFAIRGEEDTDADDGSDFNYPASGTEDQKQKIADRMRSWRMNTGG 120
Db 61 CPQCKTKYKRHKGSFAIRGEEDTDADDGSDFNYPASGNDQKQKIADRMRSWRMNTGG 120
Qy 121 SGNVCHPKYDSGEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMHMSPTGNISRR 180
Db 121 SGNVCHPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGEIPGASPDHMHMSPTGNIGRR 180
Qy 181 APFPVYNHSPNPSRFFSGISGNVANKERVYDGMKQDKGAIPTWNGTSTAPSEGRAATDI 240
Db 181 APFPVYNHSPNPSRFFSGISGNVANKERVYDGMKQDKGAIPTWNGTSTAPSEGRVGD 240
Qy 241 DASTYNNMEDALLNDETROPQSRKVPPIASSKINPYRMVIVLRLVLSIFLHRLNPNVRN 300
Db 241 DASTYNNMEDALLNDETROPQSRKVPPIASSKINPYRMVIVLRLVLSIFLHRLNPNVRN 300
Qy 301 AYPLMLLSVICIWFALSILAVDYPVKVSCVSDDGASMLTFDALAETSEFARKWVPV 360
Db 301 AYPLMLLSVICIWFALSILAVDYPVKVSCVSDDGASMLTFDALAETSEFARKWVPV 360
Qy 361 TVDPLKEPPIVTANTVLSILAVDYPVKVSCVSDDGASMLTFDALAETSEFARKWVPV 420
Db 361 TVDPMKEPPIVTANTVLSILAVDYPVKVSCVSDDGASMLTFDALAETSEFARKWVPV 420
Qy 421 KKYDTEPRAPYFQKIDYLDKQVQPSVVDKRRAMKREYEFKIRINALVSKALKVPBE 480
Db 421 KKYNTIEPRAPYFQKIDYLDKQVQPSVVDKRRAMKREYEFKIRINALVSKALKVPBE 480
Qy 481 GWIMODGTPWPGNTRDHPGMIQVFLHSGGLDTEGNELPRLVYVSREKRPQFQHHKAG 540
Db 481 GWIMODGTPWPGNTRDHPGMIQVFLHSGGLDTEGNELPRLVYVSREKRPQFQHHKAG 540
Qy 541 ANNALVRVSAVLNNGOYMLNDCDHYINNSKAVREAMCFMDPNLGPQVCYVQFPORFDG 600
Db 541 ANNALVRVSAVLNNGOYMLNDCDHYINNSKAVREAMCFMDPNLGRSVCIYVQFPORFDG 600

```
QY 601 IDNDRVANNTVFEDINLRGLDGIQGVVYVGTGCVENRTAIYGYEPPKAKKPGFLASL 660
DB 601 IDNDRVANNTVFEDINLRGLDGIQGVVYVGTGCVENRTAIYGYEPPKAKKPGFLASL 660
QY 661 CGGKKKASKSKSSDKKKSKHVDSSVPVFNLEDEEGVEGAGFDDEKSVLMSQMSLEK 720
DB 661 CGGKKKASKSKK-GSDKKKSKHVDSSVPVFNLEDEEGVEGAGFDDEKSVLMSQMSLEK 719
QY 721 RFQGSAAFAVASTLMEYGGVQPSSTPESLLKEAHIVISCGYEDKSEWTEIGWIYGSVTE 780
DB 720 RFQGSAAFAVASTLMEYGGVQPSSTPESLLKEAHIVISCGYEDKSEWTEIGWIYGSVTE 779
QY 781 ILTGFKMHARGWRSYVCMKRPAPKGSAPINLSDRLNOVLRLWALGSVEILLFSRHCPPLWYG 840
DB 780 ILTGFKMHARGWRSYVCMKRPAPKGSAPINLSDRLNOVLRLWALGSVEILLFSRHCPPLWYG 839
QY 841 YGRLKFLERFAFAYINTTIYPLTSLPLIYVYCIIPALICLLTGKFIPEISNLSIAWFIALFL 900
DB 840 YGRLKFLERFAFAYINTTIYPLTSLPLIYVYCIIPALICLLTGKFIPEISNLSIAWFIALFL 899
QY 901 SIFATGILEMRWSGVGIDEWNRNEQFWVIGGISAHLPFAVFGLLKVLGIDTNTFTVTSKA 960
DB 900 SIFATGILEMRWSGVGIDEWNRNEQFWVIGGISAHLPFAVFGLLKVLGIDTNTFTVTSKA 959
QY 961 NDEGDPFAELYMFKWTTLLIPPTTILINMVGVVAGTSYAINSGYOSWGPLFGKLPFAFW 1020
DB 960 SDEGDPFAELYMFKWTTLLIPPTTILINMVGVVAGTSYAINSGYOSWGPLFGKLPFAFW 1019
QY 1021 VIVHLVPFLKGLMGRQNRPTIIVVWAILLASIFSLWVRVDPPTTRLAGNTQTCCGINC 1080
DB 1020 VIVHLVPFLKGLMGRQNRPTIIVVWAILLASIFSLWVRVDPPTTRLAGNTQTCCGINC 1079

RESULT 3
Q9LLI6 PRELIMINARY; PRT; 1077 AA.
AC Q9LLI6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cellulose synthase-4.
GN CESA-4.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xoconostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene
  family."
RL Plant Physiol. 123:1313-1324 (2000).
DR EMBL; AF200528; AAF89964.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
DR GO; GO:0030244; P:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose_synth.
DR InterPro; IPR001841; Znf-ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00194; RING; 1.
SQ SEQUENCE 1077 AA; 120501 MW; 5F960B4AA753E2D6 CRC64;

Query Match 93.8%; Score 5421.5; DB 10; Length 1077;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 998; Conservative 46; Mismatches 33; Indels 3; Gaps 2;

QY 1 MDGDADALKSGRHGAGDVQCICADGLGTTLDGVDFTACDVCRFPVCRPCYEHKKEGTQA 60
DB 1 MEGDADGVKSGRRGGVCQICGIGVGTAGDVFAACDVCGFPVCRPCYEHKKEGTQA 60
```

```
QY 61 CLOCKTKYKHHRGSPAIRBEGDDTDADGSDENYPASGTEDEKOKIADRWRSWRMTGG 120
DB 61 CPOCKTKYKHHKGSPAIRBEGDDTDAD--SDNYLASGNEDEKOKIADRWRSWRMTGG 118
QY 121 SGNVGHKPYDGSIGIGLSKYDSGEIPRGYVPSVNSQMSGIEPCASPDHMMWSPGTNISR 180
DB 119 SGVGRPKYDGSIGELGLTKYDSGEIPRGYIPSVTNSQISGIEPCASPDHMMWSPGTNIGKR 178
QY 181 APPFYVNHSPNPREPSGSGTGNVAMKERVDMKWKQKGAIPMTNGTSPAPSEGRAATDI 240
DB 179 APPFYVNHSPNPREPSGSGTGNVAMKERVDMKWKQKGTIPMTNGTSPAPSEGRGVGDI 238
QY 241 DASTEVNMDALLNDETROPLESKVPCLASSKINPYRMVIVLRLVLSIFLHYRLTNPVRN 300
DB 239 DASTVNMEDALLNDETROPLESKVPCLASSKINPYRMVIVLRLVLSIFLHYRLTNPVRN 298
QY 301 AYPILWLLSVICETWFAISWILDQFPKWFPIINRETYLDRALRYDREGEPSQLAAVDIFVS 360
DB 299 AYPILWLLSVICETWFAISWILDQFPKWFPIINRETYLDRALRYDREGEPSQLAAVDIFVS 358
QY 361 TVDPLKEPPIVTANTVLSILAVDYPVDKVSVCYSDDGASMLTFDALAETSEFARKWVPFV 420
DB 359 TVDPMKEPPIVTANTVLSILAVDYPVDKVSVCYSDDGASMLTFDALAETSEFARKWVPFV 418
QY 421 KKYDIEBRAPEFFYFCOKIDYLDKQVPSFYKDRAMKREYEEFKIRINALVSKALKYPPE 480
DB 419 KKYNIEBRAPEWFFSQKIDYLDKQVPSFYKDRAMKREYEEFKVRVNGVLAQAQKYPPE 478
QY 481 GWMQDGTWPNGNTRDHPGMIQVFLHSGGLDTEGNEPRLVYVSRKRPFGQHHKKAG 540
DB 479 GWMQDGTWPNGNTRDHPGMIQVFLHSGGLDTEGNEPRLVYVSRKRPFGQHHKKAG 538
QY 541 AMNALVRVSAVLNGQYMLNLDCHYINNKAVERAMCFLMDPNLGPQVCYVQPPQRPFG 600
DB 539 AMNALVRVSAVLNGQYMLNLDCHYINNKAVERAMCFLMDPNLGPQVCYVQPPQRPFG 598
QY 601 IDNDRVANNTVFEDINLRGLDGIQGVVYVGTGCVENRTAIYGYEPPKAKKPGFLASL 660
DB 599 IDNDRVANNTVFEDINLRGLDGIQGVVYVGTGCVENRTAIYGYEPPKAKKPGFLASL 658
QY 661 CGGKKKASKSKSSDKKKSKHVDSSVPVFNLEDEEGVEGAGFDDEKSVLMSQMSLEK 720
DB 659 CGGKKKASKSKK-GSDKKKSKHVDSSVPVFNLEDEEGVEGAGFDDEKSVLMSQMSLEK 717
QY 721 RFQGSAAFAVASTLMEYGGVQPSSTPESLLKEAHIVISCGYEDKSEWTEIGWIYGSVTE 780
DB 718 RFQGSAAFAVASTLMEYGGVQPSSTPESLLKEAHIVISCGYEDKSEWTEIGWIYGSVTE 777
QY 781 ILTGFKMHARGWRSYVCMKRPAPKGSAPINLSDRLNOVLRLWALGSVEILLFSRHCPPLWYG 840
DB 778 ILTGFKMHARGWRSYVCMKRPAPKGSAPINLSDRLNOVLRLWALGSVEILLFSRHCPPLWYG 837
QY 841 YGRLKFLERFAFAYINTTIYPLTSLPLIYVYCIIPALICLLTGKFIPEISNLSIAWFIALFL 900
DB 838 YGRLKFLERFAFAYINTTIYPLTSLPLIYVYCIIPALICLLTGKFIPEISNLSIAWFIALFL 897
QY 901 SIFATGILEMRWSGVGIDEWNRNEQFWVIGGISAHLPFAVFGLLKVLGIDTNTFTVTSKA 960
DB 898 SIFATGILEMRWSGVGIDEWNRNEQFWVIGGISAHLPFAVFGLLKVLGIDTNTFTVTSKA 957
QY 961 NDEGDPFAELYMFKWTTLLIPPTTILINMVGVVAGTSYAINSGYOSWGPLFGKLPFAFW 1020
DB 958 SDEGDPFAELYMFKWTTLLIPPTTILINMVGVVAGTSYAINSGYOSWGPLFGKLPFAFW 1017
QY 1021 VIVHLVPFLKGLMGRQNRPTIIVVWAILLASIFSLWVRVDPPTTRLAGNTQTCCGINC 1080
DB 1018 VIVHLVPFLKGLMGRQNRPTIIVVWAILLASIFSLWVRVDPPTTRLAGNTQTCCGINC 1077

RESULT 4
Q84M43 PRELIMINARY; PRT; 1073 AA.
ID Q84M43
```

AC Q84M43;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative cellulose synthase catalytic subunit.
GN OSUNBA0059E14.5.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Van K.S., Riedmuller S.B., Uterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSUNBA0059E14 genomic sequence.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC135958; AAP21426.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; F:cellulose biosynthesis (UDP-forming) activity; IEA.
DR GO; GO:0030244; P:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose synth.
DR Pfam; PF03552; Cellulose synt. 1.
SQ SEQUENCE 1073 AA; 120681 MW; 55B677033E77A7BE CRC64;

Query Match 89.5%; Score 5169.5; DB 10; Length 1073;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 950; Conservative 67; Mismatches 54; Indels 5; Gaps 4;

QY 6 DALKSGRGAGDVCOICADGLGTLTLDGVFTACDVRCPVCYEHKKEGTQACLOCK 65
DB 2 DKAASKGQ - CHVCQCGDGVTAADGELFTACDVGFPVCPYERKKGDSQACPOCK 59
QY 66 TKYKHRGSPAIRBEGDDTDADGSDFNYPASGTEDQKQIADRMRSWRMTGSGNVG 125
DB 60 TKYKHKSGPPLIGDESDDVADADSDVNYPTSGNQDHKHAERMLTWRMNSGRDDIV 119
QY 126 HPKYSGEIGLSKYDSGEIPRGYPVPSVNSQMSGEIPGASPDHMSPTGNISRRRA - PFP 184
DB 120 HSKYDSGEIGHPKYDSGEIPRTYIPSLTHSQISGEIPGASPD - HMMSPVGNIGRRGHPPP 178
QY 185 VYNHSPNPSRPSGSGIGNVAKWERYDVGMKODKGAIPMTNCTSTAPSEGRAATDIDAST 244
DB 179 VYNHSPNPSRPSGSGIGNVAKWERYDVGMKODKGAIPMWANGTSAPSEGRGVGDIDAST 237
QY 245 EYNMEDALLNDETROPLSRKVP IASSKINPYRMVILRLVLSIFLHRLTNFVRNAYPL 304
DB 238 DYNMEDALLNDETROPLSRKVP IASSKINPYRMVILRLVLSIFLHRLTNFVRNAYPL 297
QY 305 WLLSVICSIWALSILWIDQPKWFPINRETYLDRLALRYDREGPSQAAVDIFVSTVDP 364
DB 298 WLLSVICSIWALSILWIDQPKWFPINRETYLDRLALRYDREGPSQAAVDIFVSTVDP 357
QY 365 LKEPPIVTANTVLSILAVDYPVDKYSVYSDDGASMLTFDALAESEFARKWVPVKYD 424
DB 358 MKEPPLVTANTVLSILAVDYPVDKYSVYSDDGASMLTFDALAESEFARKWVPVKYD 417
QY 425 IEPRAPEFYCQIDYLDKQVSPFKDORRAMKREYEFKIRINALVSKALVPBEGWIM 484
DB 418 IEPRAPEFYFAQIDYLDKQVSPFKDORRAMKREYEFKIRINALVSKALVPBEGWIM 477
QY 485 ODGTWPNGNTRDHPGMIQVFLHSGGLDTEGNELPRLVYVYREKRGFGHKKAGAMNA 544

DB 478 ODGTWPNGNTRDHPGMIQVFLHSGGLDTEGNELPRLVYVYREKRGFGHKKAGAMNA 537
QY 545 LVRVSAVLITNGOYMLNLDCDDHYNNKSKAVREACFLMDPNLGPVYCYVQFPORFGIDRN 604
DB 538 LVRVSAVLITNGOYMLNLDCDDHYNNKSKALREACFLMDPNLGPVYCYVQFPORFGIDRN 597
QY 605 DRYANRNTVFDDINLRGLDGIQGPVYVGTGVNFNTAIYGYEPPPIKAKKPGFLASLCGGK 664
DB 598 DRYANRNTVFDDINLRGLDGIQGPVYVGTGVNFNTAIYGYEPPPIKAKKPGFESSLCGGR 657
QY 665 KXASKSKRSDDKKKSHKVDSSVPVFNLEDEIEEGVEGAGFDDKSVLMSQMSLKRFPQ 724
DB 658 KTKKSKKSKTEKSHKSHVDSSVPVFNLEDEIEEGVEGAGFDDKSVLMSQMSLKRFPQ 717
QY 725 SAAFAVASTLMEYGVQSSPTSELKAEHAIHVISCYEDKSEWGTGIMYIGSVTDILT 784
DB 718 SSVFVASTLMEYGVQSSPTSELKAEHAIHVISCYEDKSEWGTGIMYIGSVTDILT 777
QY 785 FPMHARGRSVYCMKRPAPKGSAPINLSRLNQVLRWALGSVEILFSRHCHCPLWYGYGR 844
DB 778 FPMHARGRSVYCMKRPAPKGSAPINLSRLNQVLRWALGSVEILFSRHCHCPLWYGYGR 837
QY 845 LKFLERPAYINTIYPLTSLPLLYCILPAICLLTGKPIIMPBIINSLASIWFTALFSLIPA 904
DB 838 LKFLERPAYINTIYPLTSLPLLYCILPAICLLTGKPIIMPBIINSLASIWFTALFSLIPA 897
QY 905 TGILEMRWSGVGIDEMWRNEOPFWIGGISAHLFAVFOGLLKVLAGIDTNFTVTSKANDEE 964
DB 898 TGILEMRWSGVGIDEMWRNEOPFWIGGISAHLFAVFOGLLKVLAGIDTNFTVTSKANDEE 957
QY 965 GDFASLYMFKWTTLLIIPPTTILINMVGWAGTSYAINSGVSGWGLFGLKFFAFWVIVH 1024
DB 958 GDFASLYMFKWTTLLIIPPTTILINMVGWAGTSYAINSGVSGWGLFGLKFFAFWVIVH 1017
QY 1025 LYPFLKGLMGRQRTPTTIVWVALLASIFSLWLVRVDPFTTRLAGPNIQTGKINC 1080
DB 1018 LYPFLKGLMGRQRTPTTIVWVALLASIFSLWLVRVDPFTTRTAGPNIQTGKINC 1073

RESULT 5
Q9LLI5 PRELIMINARY; PRT; 1076 AA.
AC Q9LLI5;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Cellulose synthase-5.
GN CESA-5.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xoonostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene family.";
RL Plant Physiol. 123:1313-1324 (2000).
DR EMBL; AF200529; AAF89965.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
DR GO; GO:0030244; P:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose synth.
DR Pfam; PF03552; Cellulose synt. 1.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 1076 AA; 120724 MW; 41EE7C20EDA54F27 CRC64;

Query Match 89.0%; Score 5143; DB 10; Length 1076;
Best Local Similarity 87.6%; Pred. No. 0;


```
Db 23 DGDVPGSAKPTKSANGQVCQICGDSVGSVATGDFVACNECAFPVCRPCYEVEREGNQ 82
Qy ||| : | : | ||||| : | : ||||| : | : ||||| : | : ||||| : | : ||||| :
61 CLOCTKYKRGHSGPAIRGEEDDDADD--GSDFNYPASCTEDQKQIADRMRSWRMTG 119
Db ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
83 CPQCTRYKQKSGSPRVGDE--DEEDVDLDNEFN-----KQ----- 119
Qy 120 GSGNVGHPKYSGETGLSKYDSGEIPRGVVPVNTN--SOMSGBIPGASPDH--MMSPTGN-- 176
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
120 GSGKPEWQLQDDADLSSARHE--PHRIPLTSGQISGIBIPDASDRHSIRPTSGY 178
Qy 177 ISRAPFPVNHSPNPSREF--SGSIGNAVWVKRVDGKMKQDKGAIPMTNGTSIAPSEGR 235
Db : ||| : : | : | : | : | : | : | : | : | : | : | : | : | : | :
179 VDPSPVPEV--RIVDPKDLNSGLNSVDKERVESWRVKQDKNMQVNTN---KYPEAR 232
Qy 236 AATDIDASTENMEDALLNDETROPKVPITASSKINPVMVIVLRVLVLSIFLHYBLT 295
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
233 GG-DME-GTSGNGEXKQWMDVDAARLPLSRIVPISSNQLNLYRVILRLIILCFYQYRVS 290
Qy 296 NPVRNAYPLWLLSVTCEIWFALSWILDQFPKWFPPINRETYDLRLALRYDREGEPSQAAV 355
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
291 HPVRDAYGLWLVSVCEWFWALSWILDQFPKWPINRETYDLRLALRYDREGEPSQAPI 350
Qy 356 DIFVSTVDPLKEPPIVTANTVLSILAVDPVKVSCYVSDDGASMLTFDALAETSEFARK 415
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
351 DVFVSTVDPLKEPPIVTANTVLSILSVDPVDPVKVSCYVSDDGASMLTFESLSETAEFARK 410
Qy 416 WVPVKYDIIPRAPEFEFCOKIDVLDKQVPSFVKDRAMKREYEEFKIRINALVSKAL 475
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
411 WVPCKKHNIIPRAPEFYFAQIDVLDKQIOPSFVKERRAMKREYEEFKIRINALVAKAQ 470
Qy 476 KVPBEGWIMQGTWPWGNNTDHPGMIQVFLGHSGGLDTEGNELPRLVYVSRKRPQGH 535
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
471 KVPBEGWIMQGTWPWGNNTDHPGMIQVFLGHSGGLDTEGNELPRLVYVSRKRPQGH 530
Qy 536 HKKAGMAMALVRVSAVLNNGOYMLNLDCHYINNSKAVREAMCFMLDPLNGFQVCYQFP 595
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
531 HKKAGMAMALVRVSAVLNNGOYMLNLDCHYINNSKAVREAMCFMLDPLNGFQVCYQFP 590
Qy 596 QRPDGDIDRNDYANRNTVFEDINRLGLDGIQPVYVGTGCVFNRTAIYGYBPPTKAK--K 653
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
591 QRPDGDIDRNDYANRNTVFEDINRLGLDGIQPVYVGTGCVFNRTAIYGYBPPTKAK--K 650
Qy 654 PGFLASLGGKKKSKSKSSDKKSKKHVDSSVVPFNLEDIEEGVEGAGFDPDEKSKVM 713
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
651 PNVIKSCGGRK--KKNKSYMDSOSRIMKRTESSAPFNEMEDIEEGIE--GYEDERSVLM 707
Qy 714 SQMLEKRFQGSAAFAVASTLMYEGVQSPSTPESLLKEAIIHVISCGYEDKSEWGTGWI 773
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
708 SQRKLEKRFQGSPIFIASFTFTQGGIPPTNPASILLKEAIIHVISCGYEDKTEWKGWI 767
Qy 774 YGSVTEDLITGFKHARGSVYCMKRPAPKGSAPINLSRLNQLVRLWALGVSVEILFSR 833
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
768 YGSVTEDLITGFKHARGSVYCMKRPAPKGSAPINLSRLNQLVRLWALGVSVEILFSR 827
Qy 834 HCLPMWYGGRLKPLERFAYINTIYPLTSLPDLVYICLPAICLLTGKFMPEISNLASI 893
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
828 HCPIMWYGGRLKPLERFAYINTIYPLTSLPDLVYICLPAICLLTGKFMPEISNLASI 887
Qy 894 WFTALFISIFATGILENRWVGSDIEWNRNEQFVWIGISAHFLFAVFGQLKVLGADITN 953
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
888 FFILLFASIFATGILELRWSVGIEDWNRNEQFVWIGISAHFLFAVFGQLKVLGADITN 947
Qy 954 FTVTSKANDREGDFAELIMKWTLLIPPTTILLINMGVVGAGTSYAINSGYSQWGLPFG 1013
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
948 FTVTSKASDEGDFAELIVFKWTSLLIPPTTIVLINVGWVAGISYAINSGYSQWGLPFG 1007
Qy 1014 KLFFAFVIVVHLYPFLKGLMGQRNRTTIVIVWAVILLASIFSLIWRVDPDP---TTLAG 1070
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1008 KLFFSIWVILHLYPFLKGLMGQRNRTTIVIVWAVILLASIFSLIWRVDPDP---TTLAG 1070
Qy 1071 PNQTCTGICNG 1080
: ||| :
```

```
Db 1068 --LGQCGVNC 1075
RESULT 11
O48946
ID O48946 PRELIMINARY; PRT; 1081 AA.
AC O48946;
DC 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Cellulose synthase catalytic subunit (RSM1).
GN RSM1 OR F8B4.110 OR AT4G32410.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=98111412; PubMed=9445479;
RA Arioli T., Peng L., Betzner A.S., Burn J., Wittke W., Herth W.,
RA Camilleri C., Hofte H., Plazinski J., Birch R., Cork A., Glover J.,
RA Redmond J., Williamson R.E.;
RT "Molecular analysis of cellulose biosynthesis in Arabidopsis.";
RL Science 279:717-720(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Bevan M., Terry N., Ardiles W., Buyshaert C., Dasseville R.,
RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,
RA Villarroel R., Gielen J., Van Montagu M., Hoheisel J., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Terry N., Ardiles W., Buyshaert C., Dasseville R., De Clerck R.,
RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villarroel R.,
RA Gielen J., Van Montagu M., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Yamada K., Dale J.M., Hsuan V.W., Onodera C.S., Quach H., Toriumi M.,
RA Wong C., Wu H.C., Yu G., Yuan S., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Arabidopsis Full length cDNA Clones.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [7]
RP EMBL; AF027172; AAC39334.1; -.
RN [8]
RP EMBL; AL034567; CAA22568.1; -.
RN [9]
RP EMBL; AL161581; CAB79958.1; -.
RN [10]
RP EMBL; BT008654; AAP40467.1; -.
RN [11]
RP FIR; T05351; T05351.
RN [12]
RP GO; GO:0016020; C-membrane; IEA.
RN [13]
RP GO; GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
RN [14]
RP GO; GO:0030244; P:cellulose biosynthesis; IEA.
RN [15]
RP InterPro; IPR005150; Cellulose_synth.
RN [16]
RP InterPro; IPR001841; Znf_ring.
RN [17]
RP Pfam; PF03552; Cellulose_synt; 1.
RN [18]
RP SMART; SM00184; RING; 1.
RN [19]
SQ SEQUENCE 1081 AA; 122236 MW; BDEB5D9DEE334D59 CRC64;
Query Match 68.6%; Score 3963.5; DB 10; Length 1081;
Best Local Similarity 68.0%; Pred. No. 3e-311;
```


Db 473 NALVAKQKVPEEGWTMDGSPWPGNNVRDHPGMIQVFLGSGGRDVEGNELPLRVVYSR 532
Qy 528 EKRPQFOHHKAGAMNALVRVSALVTGQYMLNLDCHYINNSKAVREAMCFMDPNLGP 587
Db 533 EKRPYNNHKKAGAMNALVRVSALVSNAAVYLLNLDCHYINNSKAIKEAMCFMMDPLVGG 592
Qy 588 QVCYVQPPQRPDGDIDRNDRYANRNVTFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEP 647
Db 593 KVCYVQPPQRPDGDIDRNDRYANRNVTFDINMKGLDGIQGPVYVGTGCVFRQALYGYDA 652
Qy 648 PIKAKKP-----GFLASLGGKKKASKSKRSDDKKKS--NKHVDSVVPFNL 695
Db 653 P-KTKKPPSRTCNCPKWCWCLSCCNRKNKXKTKTKPKTEKKKLFKKAENPSPAYALGE 711
Qy 696 IEQVEGAGDEDEKSVLMSQMSLEKRPCOSAAFVASTILMEYGGVQPSSTPESILKEAIVH 755
Db 712 IDEGAPGA--DIEKAGIVNQKLEKFGQSSVFVASTLLENGGTLKSPASLKEAIVH 769
Qy 756 ISCYEDKSEWGTETIGYISVTEIDITLGFQNHARGWRSVYCMKRPFAFGSAPINLSDR 815
Db 770 ISCYEDKTDWKEIGMIYISITIEDITLGFQMHCHGWSYICIPKRPFAFGSAPINLSDR 829
Qy 816 LNOVLRWALGSVEILFRSHCPCLWYGYGRKLFERFAYINTIYPLTSLPLVYCIILPAI 875
Db 830 LHQVLRWALGSVEIFFSKHCPCLWYGYGGKLFERFSYINSIVYVWTSIPLLAYCTILPAI 889
Qy 876 CLLTGKTEIMPEISNLASIFIALFSLFATGILEMRWSGVGIDEWNEQFVWIGTSAH 935
Db 890 CLLTGKTEIPELTNVASIFWMAFICISVGTGILEMRWSGVGIDEWNEQFVWIGTSAH 949
Qy 936 LFVAFQGLLVKLAGIDNTFTVTSKANDBEQDFAELYMFKWTLLIPPTTILINMVGVA 995
Db 950 LFVAFQGLLVKAVAGIDTSFVTSKAGDDE-EFSLEYTFKWTLLIPPTTILINMVGVA 1008
Qy 996 GTSVAINSGYQSGWGLFGKLPFAFWIVHLYPLFKLGMGRQNRPTTIVIVAVLLASIFS 1055
Db 1009 GISNAINNGYESWGFLFGKLPFAFWIVHLYPLFKLGLVGRQNRPTTIVIVMSILLASIFS 1068
Qy 1056 LLWVRVDPFTTRLAGNPVQTCGNC 1080
Db 1069 LLWVRVDPFLAKNGPLLEBGGDUC 1093

RESULT 13
Q851L8 PRELIMINARY; PRT; 1092 AA.
AC Q851L8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cellulose synthase.
GN CSJNEA0042109.26.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Teitrit T., Kim M.M., Bera J.J., Jin S.S.,
RA Radrosh D.W., Tallon L.J., Koo H., Zismann V., Heiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNBa0042109 genomic sequence.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC04487; AAC41140.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
DR GO; GO:0030244; P:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose_synth.

DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
SQ SEQUENCE 1092 AA; 123395 MW; A682DOAA434BE24D0 CRC64;
Query Match 66.3%; Score 3833; DB 10; Length 1092;
Best Local Similarity 65.1%; Pred. No. 1.1e-300;
Matches 719; Conservative 147; Mismatches 177; Indels 62; Gaps 17;
Qy 2 DGDADALKSRHGAGDVQCICADGLGTLTLDGVFTACDVCRCFPVPCRCYEHKEGTCAC 61
Db 23 DGEKGP-KPVKHTGQVCQICGDDVGLTPDGEPPVACNECAFPVCRDCYEREREGTQC 81
Qy 62 LOCKTKYKRRGSPAIRGEGDDTDADD--GSDFNYPASGTEDQKQKADMRWRMNTGG 120
Db 82 PQCKTRFKRLKGCARVPGDE--EEDVDLLENFNM--RDKTDQYVAESMLHGMYSGR 137
Qy 121 SGNV-GHPKYDSGEBGLSKYDSGEBIPRGYPVSTNSQMSGEBIPGASPDHMMST--GN 176
Db 138 GGDLDGVQPHQ-----PIPNVPLLTNGEMADDIP---PEQHALVPSFMGG 181
Qy 177 ISSRAPPPYVN-----HSPNPSREFSG-SIGNVAWKERVDMKQDKGAIPTMGTS 228
Db 182 GKRIHPLPYADENLPVQPRSMPSKDLAAVGYGSAVKERMESWKQKQERLHQMRDQ-- 239
Qy 229 IAPSEGRAATDIDASTEYNMEDALLNDETQPLSRKVPDIASSKINPYRMVTLRLVLSI 288
Db 240 -----GGKMDGDDGDD---ADLPMDARQPLSRKIPISSSLVNRYMIIIRLVVLF 290
Qy 289 FLHRLNPNVRNAVPIWLLSVICEIFALSWILDQFPKWPINRETYLDRALRYDREGE 348
Db 291 FFHYRVWHPVDFALMLSVICEIFAMSWILDQFPKWPFIETRYLDRULRFDKEGQ 350
Qy 349 PSQALAAVDIFVSTVDPKPEPIVANTVLSILAVDYPDKVSCVSDGASMLTFDALAE 408
Db 351 QSQALAVDFVSTVDPKPEPLVANTVLSILAVDYPDKVSCVSDGASMLTFEALSE 410
Qy 409 TSFAPKWWVFKYDIEPRAPFYYFOKIDYLDKQVQPSFVDRRRAMKREYEFKIRIN 468
Db 411 TSFAPKWWVFKYDIEPRAPFYYFOKIDYLDKQVAPNFPVRRERRAMKREYEFKIRIN 470
Qy 469 ALYSKALKVPEEGWIMODGTPWPGNNTRDHPGMIQVFLGSGGDLTSGNELPLRVVYSRE 528
Db 471 ALVAKQKVPEEGWTMDGSPWPGNNVRDHPGMIQVFLGSGGHDVEGNELPLRVVYSRE 530
Qy 529 KRPFQHHKKAGAMNALVRVSALVTGQYMLNLDCHYINNSKAVREAMCFMDPNLGP 588
Db 531 KRPGYNHKKAGAMNALVRVSALVTNAPYMLNLDCHYINNSKAIKEAMCFMMDPLVGGK 590
Qy 589 VCYVQPPQRPDGDIDRNDRYANRNVTFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPP 648
Db 591 VCYVQPPQRPDGDIDRNDRYANRNVTFDINMKGLDGIQGPVYVGTGCVFRQALYGYDAP 650
Qy 649 IKAKKQFGLASLC-----GGKKKASKSKRSDDKKKS--NKHVDSVVPFNL 695
Db 651 -KSKKPPSRTCNCPKWCWCLSCCNRKNKXKTKTKPKTEKKKLFKKAENQSPAYALGE 709
Qy 696 IEQVEGAGDDEKSVLMSQMSLEKRPCOSAAFVASTILMEYGGVQPSSTPESILKEAIVH 755
Db 710 IDEGAPGA--ENEKAGIVNQKLEKFGQSSVFVASTLLENGGTLKSPASLKEAIVH 767
Qy 756 ISCYEDKSEWGTETIGYISVTEIDITLGFQMHCHGWSYICIPKRPFAFGSAPINLSDR 815
Db 768 ISCYEDKTDWKEIGMIYISVTEIDITLGFQMHCHGWSYICIPKRAAFKGSAPINLSDR 827
Qy 816 LNOVLRWALGSVEILFRSHCPCLWYGYGRKLFERFAYINTIYPLTSLPLVYCIILPAI 875
Db 828 LHQVLRWALGSVEIFFSKHCPCLWYGYGGKLFERFSYINSIVYVWTSIPLLAYCTILPAI 887
Qy 876 CLLTGKTEIMPEISNLASIFIALFSLFATGILEMRWSGVGIDEWNEQFVWIGTSAH 935
Db 888 CLLTGKTEIPELTNVASIFWMAFICISVGTGILEMRWSGVGIDEWNEQFVWIGTSAH 947

```
QY 936 LFAVFQGLLVLAGIDTNTFTVTSKANDEBGFALYMEKWTLLIPPTTILINNVGVVA 995
Db 948 LFAVFQGLLVAGIDTFTVTSKGDDE-EFSELYTFKWTLLIPPTTLLLNFIGVVA 1006
QY 996 GTSVAINSGYSWGLPFGKLFPAFVWVHLVPLKGLMGRQNRTPTIIVWAVLLASIFS 1055
Db 1007 GVSNAINGYESWGLPFGKLFPAFVWVHLVPLKGLVGRQNRTPTIIVWVLSILLASIFS 1066
QY 1056 LLWVRVDPFTTLAGPNIQTGGINC 1080
Db 1067 LLWVRIDPFLAKNDGPLEEBCGLDC 1091

RESULT 14
Q9LLI14 PRELIMINARY; PRT; 1059 AA.
AC Q9LLI14;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
D3 Cellulose synthase-6.
GN CESA-6.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xocoostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene
RT family";
RL Plant Physiol. 123:1313-1324 (2000).
DR EMBL; AF200530; AAF89966.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; P:cellulose synthase (UDP-forming) activity; IEA.
DR GO; GO:0030244; P:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose_synth.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 1059 AA; 120044 MW; 289DA26B25232249 CRC64;

Query Match 66.1%; Score 3820; DB 10; Length 1059;
Best Local Similarity 65.1%; Pred. No. 1.2e-299;
Matches 714; Conservative 140; Mismatches 168; Indels 74; Gaps 16;

QY 16 GDVCOICADGLGTLDGVDFTACDVCPPVPCVYEHKEGTQACLOCKTKYKGRGSP 75
Db 6 GQVCQIGDDVGRNDFGPFVACNECAPFICRDCVEYERREGTQPCQKTRFKFKGCA 65
QY 76 AIRGEGDDTDADGSDFNYPASGTEDQKQIADRMRSWRMNTGGSNVG-----HPKY 129
Db 66 RVPGEDEEDGVLDLENEFW----SKHDSQYLAESMLHAHMSYGRGADLDGVQPPEHP-- 120
QY 130 DSGEITGLSKYDGETPRGVVSVTMSOMGETPGASPDDHMMSP-----GNISRRAPFPYV 186
Db 121 -----IPN-----VPLITNGQWVDIP---PQHALVPSFVGSGGRIIHLPLPYA 161
QY 187 N-----HSPNPSREFSG-SIGNVAWKERVQGMKQDKGAIPMTNGTSIAPSEGRAAT 238
Db 162 DNLVPQPSMPSDKDLAAYGVSVAWKERMESWQKQER--MHQTR-----NDGGDD 213
QY 239 DIDASTEYNMEDALLNDETQPLSRKVPITASKINPYRMVIVLRLVLSIFLHYRLTNPV 298
Db 214 GDGA-----DLPLMDEARQPLSRKIPLPSSQINPYRMIIIRLVLCFFFFHYRVMHPV 266
QY 299 RNAPYPLWLSVCEIWFALSILDQPKWFPINRETYLDRLALRYDREGESQLAADVIF 358
Db 267 PDAFALWLSVCEIWFAMSWILDQPKWFPFERETYLRLSLRFDKEGHPSQLAPVDFF 326
```

```
QY 359 VSTDPLKEPPIVTANTVLSILAVDPVDKVS CVYSDDGASMLTFDALAETSEFARKWVP 418
Db 327 VSTDPLKEPPIVTANTVLSILSVDPVDKVS CVYSDDGAAMLTEALSETSEFARKWVP 386
QY 419 FVKKYDIDPRAPEFFCQKIDYLDKQVOPSVFKDRAMKREYEEFKIRINALVSKALVP 478
Db 387 FCKRYSLEPRAPEFWYFQKIDYLDKQVAPNFVRERRAMKREYEEFKIRINALVAKAQVP 446
QY 479 BEGTMQDGTWPNGNTRDHPQMIOVFLGSHSGGLDTEGNEPRLVYVSRKRPQPOHKK 538
Db 447 BEGTMQDGTWPNGNVRDHPQMIOVFLGSGGHDVEGNEPRLVYVSRKRPQGNHHKK 506
QY 539 AGAMNALVRVSAVLTNGQYMLNLDCHYINNSKAVREAMCFMLDPNLGPQVCYVQPPQRF 598
Db 507 AGAMNALVRVSAVLTNAPYLLNLDCHYINNSKAIKEAMCFMMDPLLGGKVCYVQPPQRF 566
QY 599 DGIDRNDRYANRNTVFFDINRLGLDGIQGVVVGTCVFNRTAIYGYBPPIKAKXP---- 654
Db 567 DGIDRHDYANRNVVFFDINMKGLDGIQGVVVGTCVFRQALYGYDAP-KTKKPPGRT 625
QY 655 -----GFLASLCGKKKASKSKKSSDK-----KSKNHVDSSVPVFNLEDIEEGVEGAG 704
Db 626 CNCWPKWICCCCFGNRTKTKTKSKPKFEKIKLKKENQAPAYALGEIDEAAPGA- 684
QY 705 FDDEKSVLMSQMSLEKRFQSQAAVASTLMYEGVGPQSTPESLLKEAIIHVISCYEDKS 764
Db 685 -ENEXASIVNQKLEKFKQSSVFAVSTLLENGGTLKGSASPASILLKEAIIHVISCYEDKT 743
QY 765 EWGTETGHWYSVTEDILTGFQKHARGHRSVYCMKRPAPKGSAPINLSDRNLNOVLWAL 824
Db 744 GWGKDIGHYGSVTEDILTGFQKHCHGRWSYICIPKRAAFKGSAPINLSDRHQVLRWAL 803
QY 825 GSVELFSRKHCPWYGYGRLKFLERFAYINTIYPLTSLPLLYCILPAICLLTGKFTM 884
Db 804 GSIEILFSNHCPWYGYGGLKFLERFSYINSIVPWTISIPLLACTLPAICLLTGKFT 863
QY 885 PEISNLASIWFIALFSPATGILEMRWSGVGIDEMWRNEQFWITGGISAHLFAVFGGL 944
Db 864 PELNNVASLWFMSLFICIPATSIEMRWSGVGIDDMWRNEQFWITGGVSSHLFAVFGGL 923
QY 945 KVLGIDNFTVTSKANDEBGFALYMEKWTLLIPPTTILINNVGVVAGTSVAINSG 1004
Db 924 KVIAGVDTSTVTSKGDDE-EFSELYTFKWTLLIPPTTLLLNFIGVAGTSVAINSG 982
QY 1005 YOSWGLPFGKLFPAFVWVHLVPLKGLMGRQNRTPTIIVWAVLLASIFSLLWVRVDPF 1064
Db 983 YESWGLPFGKLFPAFVWVHLVPLKGLVGRQNRTPTIIVWVLSILLASIFSLLWVRIDPF 1042
QY 1065 TTRLAGPNIQTGGINC 1080
Db 1043 LAKDDGPILLECGCLDC 1058

RESULT 15
Q9LLI13 PRELIMINARY; PRT; 1086 AA.
AC Q9LLI13;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Cellulose synthase-7.
GN CESA-7.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xocoostle-Cazares B., Delmer D.P.;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 22, 2004, 17:17:01 ; Search time 6177 Seconds
(without alignments)
5221.168 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MDGDADALKSRHGAGDVQC.....VDPFTRLAGNIQTGINC 1080

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xih
-Q=/cgn2_1/USPRO-epool/US0900237/runat_18082004_081514_14648/app.query.fasta_1.1223
-DB=EST -QFMT=fastcap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0900237@cgn_1_4365@runat_18082004_081514_14648 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=10 -XGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPOP=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5421.5	93.8	3763	11	AY108113	AY108113 Zea mays
2	4937.5	85.5	3728	11	AY112236	AY112236 Zea mays
3	4805	83.2	3696	11	AY110079	AY110079 Zea mays
4	3902	66.1	3898	11	AY110415	AY110415 Zea mays
5	3817	66.1	3783	11	AY104730	AY104730 Zea mays
6	3806	65.9	3897	11	AY103655	AY103655 Zea mays
7	3805	65.9	3788	11	AY103701	AY103701 Zea mays
8	3525	61.0	2872	11	AY104236	AY104236 Zea mays
9	2507	43.4	1874	14	CD726831	CD726831 EST027 Cu
10	1575	27.3	916	14	CK272603	CK272603 EST18681
11	1542	26.7	924	14	CF513822	CF513822 CABud0007
12	1530.5	26.5	977	14	CK271294	CK271294 EST17372
13	1503.5	26.0	958	14	CK274934	CK274934 EST721012
14	1463.5	25.3	978	12	BG321258	BG321258 Zm04.03f1
15	1430	24.7	953	14	CD439980	CD439980 EL01N0531
16	1429.5	24.7	874	14	CD433620	CD433620 EL01N0313
17	1422	24.6	865	14	CD440398	CD440398 EL01N0554
18	1418.5	24.6	1189	11	AY107656	AY107656 Zea mays
19	1407	24.4	880	12	BM816138	BM816138 HC109R12
20	1390.5	24.1	925	14	CF213555	CF213555 CGF100079
21	1376	23.8	875	14	CF514581	CF514581 CABud0005
22	1359	23.5	1014	29	CG203358	CG203358 POKCB90TB
23	1356.5	23.2	966	14	CA247044	CA247044 SCCAM209
24	1341	23.2	913	14	BG368813	BG368813 HVSME1002
25	1321.5	22.9	870	12	BG368813	BG368813 HVSME1002
26	1314.5	22.8	1022	28	BZ827769	BZ827769 PUF0107TB
27	1313	22.7	807	12	B1311236	B1311236 EST531298
28	1310.5	22.7	960	9	AB049602	AB049602 AB049602
29	1307	22.6	762	13	BQ802778	BQ802778 WHE2829 H
30	1306	22.6	781	14	CF447738	CF447738 EST684083
31	1294.5	22.4	837	14	CF512600	CF512600 CABud0003
32	1290	22.3	809	14	CF445367	CF445367 EST681712
33	1289.5	22.3	905	13	CA159025	CA159025 SCEZR2312
34	1289	22.3	840	13	CA151353	CA151353 SCJFR2200
35	1286	22.3	810	14	CF444535	CF444535 EST680880
36	1282	22.2	750	10	BF624748	BF624748 HVSMEa001
37	1274	22.0	1042	14	CA239344	CA239344 SCBFFL507
38	1273	22.0	851	14	CB892280	CB892280 EST649249
39	1271	22.0	1133	14	CK208799	CK208799 FGAS02052
40	1268	21.9	807	14	CA918393	CA918393 EST642540
41	1264	21.9	821	29	CG383072	CG383072 OG1BH95TV
42	1261	21.8	741	10	BF460078	BF460078 072G01 Ma
43	1259	21.8	827	14	CF437366	CF437366 EST673711
44	1256	21.7	701	14	CD930401	CD930401 GR45.111C
45	1251.5	21.7	816	14	CB097089	CB097089 AF53-Rpf_

ALIGNMENTS

RESULT 1
AY108113
LOCUS AY108113
DEFINITION Zea mays PC0126465 mRNA sequence.
ACCESSION AY108113
VERSION AY108113.1 GI:21211191
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3763)

QY	521	ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly	540	QY	881	LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu	900
Db	1892	CGTTGGTCTATGTTCTTCGTGAAGAAGCTCTCGATTCCAGCATCACAGAAGAGCTGGT	1951	Db	2969	AAGTTCATCATTTCCAGAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTCTTCATC	3028
QY	541	AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn	560	QY	901	SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr	920
Db	1952	GCCATGAATGCTCTTGTTCTGTGCTCAGCTGTGCTTACCAATGGACAATACATGTTGAAT	2011	Db	3029	TCGATCTTCGCCACGGGCATCTCGAGATGAGTGGAGGGGGTGGGATCGACAGATGG	3088
QY	561	LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu	580	QY	921	TyrArgAsnGluGlnPheTyrValIleGlyIleSerAlaHisLeuPheAlaValPhe	940
Db	2012	CTTGATTGTGATCACTACATTAACAACAGTAGGCTCTCAGGGAAGCTATGTGCTTCCTT	2071	Db	3089	TGGAGGAACGAGCAGTTCCTGGGTGATCGGGGGCATCTCCGCGCACCTCTTCGCCGTGTC	3148
QY	581	MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly	600	QY	941	GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla	960
Db	2072	ATGGACCTTAACCTTAGGAGAGGTGTGCTACGTCCAGTTCCTCCAGAGATTCGATGGC	2131	Db	3149	CAGGGCTGCTCAAGAGTGTGGCGGCATCGACCAACTTCACCGTCACCTCCAGGCG	3208
QY	601	IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg	620	QY	961	AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeuIle	980
Db	2132	ATTGACAGGAATGATCGATATGCCAACAGGAACACCGGTGTTTTCGATATTAACTTGAGA	2191	Db	3209	TCGGACGAGGACGGCGACTTCGGGAGCTGTACATGTTCAAGTGGACGACGCTCTGATC	3268
QY	621	GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr	640	QY	981	ProProThrThrIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyrAla	1000
Db	2192	GGTCTTGATGGCATCCAGGACCAAGTTTATGTGGAACCTGCTGTGTTTCAACCGAACA	2251	Db	3269	CGCCCCACCATCTGATCATCAACTGGTCGCGTCTCGCGGCATCTCTCTACGCC	3328
QY	641	AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu	660	QY	1001	IleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPhePheAlaPheTyr	1020
Db	2252	GCTCTATGTTATGAGCCCCCAATTAAGCAGAAGAGGGTGGTTCCTTGTGATCACTA	2311	Db	3329	ATCAACAGCGGATACCACTGCTGGGGCCGCTCTTCGCAAGCTCTTCTTCGCTCTCG	3388
QY	661	CysGlyGlyLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer	680	QY	1021	ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro	1040
Db	2312	TGTGGCGGTAGGAAGAGGACCAATCAAGAAG---GGCTCGGACAAGAAGATCG	2368	Db	3389	GTCATCGTCCACTGTACCCGTTCTCAAGGGCTCATGGGCGAGCAGAACCCGACCCG	3448
QY	681	AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal	700	QY	1041	ThrIleValIleValTyrAlaValLeuLeuAlaSerIlePheSerLeuLeuTyrValArg	1060
Db	2369	CAGAAGCATGTGGACAGTTCTGTGCCAGTATTCAACTTGAAGATATAGAGGAGGATT	2428	Db	3449	ACCATCGTCGTCGTGGGCCATCTGTCGGCGTCCAATCTTCTCTGCTGGTTCGC	3508
QY	701	GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys	720	QY	1061	ValAspProPheThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys	1080
Db	2429	GAAGCGCTGATTTGACGACGAGAAATCACTTTATGTCTCAAAATGAGCTGGAGAAG	2488	Db	3509	ATCGACCCCTTCCACCCGCTCATCTGGCCCGATATCCAGACGTGTGGCATCAACTGC	3568
QY	721	ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro	740	RESULT 2			
Db	2489	AGATTTCGCCAGTCGCCAGGTTTGTTCCTCCACTCTGATGGAGTATGGTGTTCCT	2548	LOCUS	AY112236	3728 bp	mRNA linear HTC 17-OCT-2002
QY	741	GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr	760	DEFINITION	Zea mays CL1160_1	mRNA sequence.	
Db	2549	CAGTCCCGCAACTCCGAGTCTCTTCTGAAGAAGCTATCCATGTTATAAGCTGTGGCTAT	2608	ACCESSION	AY112236		
QY	761	GluAspLysSerGluTyrGlyThrGluIleGlyTyrPheTyrGlySerValThrGluAsp	780	VERSION	AY112236.1	GI:21216826	
Db	2609	GAGGACAAGACTGAATGGGAACCTGAGATCGGTGGATCTACGTTCTGTGACAGAAGAC	2668	KEYWORDS	HTC.		
QY	781	IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys	800	SOURCE	Zea mays		
Db	2669	ATTCTACCGGATTCAGAGATCACCGCGAGGCTGGCGTCTACTGTCATGCCCAAG	2728	ORGANISM	Zea mays		
QY	801	ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu	820	REFERENCE	1 (bases 1 to 3728)		
Db	2729	CGGCGACCTTTCAGGGGTCTGCCCCCATCAATCTTTCGGACCGCTCTGAACAGGTGCTC	2788	JOURNAL	Unpublished (2002)		
QY	821	ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGly	840	AUTHORS	Coe,E.H.		
Db	2789	CGTGGGCTCTTGGGTCTGGAGATCTCTTCAGCGCGCACTGCCCTCTGTGTCACGGC	2848	JOURNAL	Submitted (25-APR-2002)	Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA	
QY	841	TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro	860	COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSI, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.		
Db	2849	TACGAGGGCGGCTCAAGTTCCTGGAGAGATTCCGCTACATCAACACCATCTACCCG	2908	FEATURES	Location/Qualifiers		
QY	861	LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly	880	SOURCE	1..3728	/organism="Zea mays"	
Db	2909	CTACGTCCTACCGCTTCTCATCTACTGATCTCTGCCGCCCATCTGTCTGCTCACCGGA	2968				

```

/mol_type="mRNA"
/db_xref="MaizeDB:630049"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by Dupont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 0 Length: 3728
Score: 4937.50 Matches: 918
Percent Similarity: 89.07% Conservative: 44
Best Local Similarity: 85.00% Mismatches: 117
Query Match: 85.45% Indels: 1
DB: 11 Gaps: 1

US-09-900-237A-30 (1-1080) x AY112236 (1-3728)
QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
238 ATGGAGGCGCAGCGGCGCGCTGAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTGCCAG 297
QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
298 ATCTCGCGCGATGGCGTGCGCACTACGGCGGAGGAGAGCGTCTTCACCGCCTGCGACGTC 357
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgGlyGlyThrGlnAla 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
358 TCGCGGTTCGCGGTGCGCGCCCTCTACGAGTACGAGCGCAAGAGCGCACACAGCG 417
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
418 TGNNNNNAGTGCNNNNNCAAGTACAAGCGCCACACAGGGGAGTCCAGCGATCCGAGGGAG 477
QY 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
478 GAAGGAGACGATACTGATCGCATGATGCTAGCGACTTCAACTACCTCCTCATCTGGCAAT 537
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
538 GACGACCGAAGCAGAGAAGATTGTGACAGATCGCGAGCTGGCGCATGAATGCTNNNNNC 597
QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
598 AGCGGGGATGTTGGCGCGCCCAAGTATGACAGTGTGTGAGATCGGGCTTACCAAGTACGAC 657
QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
658 AGTGTGTGATCCCTCGGGGATACATCCNNNNNNNNNNNNNNNNNNNNNNNNNNATTTTCGGGAGAA 717
QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
718 ATCCCTCGGTGCTTCCCTGACCATATATATGATGCTCTTACTCGGGAACATTCGACGGCGC 777
QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
778 GCCCCATTCCTTATATGAATCATTCATCAATCCGTCGAGGGAATTCCTCTGTGTGCGTT 837
QY 201 GlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAla 220
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
838 GGGAAATGTTCCCTGGAAAGAGAGGGTGTGATGGCTGGAAATGAAGCAGGACAAGGGAACA 897
QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
898 ATTCCCATGACGAATGGCAACAGCATTCCTCTCTGAGGGCGCGGGTGTGTGTGATATT 957
QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

DB 958 GATGCATCAACTGATTACACATGGAAGATGCTTTATTAAACGATGAACACTGCCAGCCT 1017
QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1018 CTATCTAGAAAGTTCACCTTCCTTCCAGGATAAATCCATACAGATGTCATTGG 1077
QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1078 CTACATGATTGTTCTAAGCATCTTCCTGCACTACCGGATCACAATCTCTGCGTAAT 1137
QY 301 AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle 320
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1138 GCATACCCACTGTGCTTCTATCTTATATGTGAGATCTGGTTCCTCTTTCTGGATA 1197
QY 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1198 TTGGATCAGTTTCCAAAGTGTTCCTCAATCAACCGGAGACTTACTTGTATAGATCTGCA 1257
QY 341 LeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1258 TTAAGGTATGACCGGAAAGGTGAGCCATCTCAAGTTGGCTGTCTGTGACANNNNNGTCAGT 1317
QY 361 ThrValAspProLeuLysGluProIleValThrAlaAsnThrValLeuSerIleLeu 380
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1318 ACTGTGACCCCAATGAAGAGCTCTCTGTGTCACCTGCCAAATACCGTGTCTATCCATCTC 1377
QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1378 GCTGTGGACTATCTCTGTGATGAAGTCTCTGTCTATGTATCTGTATGATGAGCTGCTATG 1437
QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1438 CTGACATTTGATGCTAGCTAGACTTCAGAGATTGTGTAGAAAAATGGTGGCTTGTGT 1497
QY 421 LysLysTyrAspIleGluProGluPheTyrPheCysGlnLysIleAspTyr 440
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1498 AAGAAGTACAACTTGAACCTAGAGCTCTGTAATGGTACTTCTCCAGAAAAATTGATTAC 1557
QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaAlaMetLysArgGluTyr 460
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1558 TTGAAGGACAAAGTGCACCTTCATTTGTTAAAGCCGCGGNNNNNNNNNNNNNNNNNN 1617
QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1618 NNNNNNNCAAATAATTAGGCTAAATGGCTTGTGTAGGCAACNNNNNGTCCCTGAGGAA 1677
QY 481 GlyTrpIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisProGly 500
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1678 GGATGGATCATGCAAGATGGCACACCATGGCCAGGAAACAAATACAGGGACCACTCTGGA 1737
QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1738 ATGATTGAGTTTCTTGTGTACAGTGTGTCTTGTATCTGAGGGTAAATGAGCTACCC 1797
QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly 540
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1798 CGTTTGTCTATGTTTCTCTGNNNNNGTCTCTGGATTCAGCATCACAGAAAGTGGT 1857
QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1858 GCCATGAATGCTCTTGTCCGCGTCTCAGCTGTGCTTACCAATGGACAAATATGTTGAAT 1917
QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1918 CTTGATTGTGATCATTACATCAACACAGTAAAGCTCTCAGGGAAGCTATGTGCTTCTT 1977
QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1978 ATGCATCTCAACCTAGGAAGAGTGTCTGTATGTTTCAAGTTCCTCCAGAGGTTTCGATGGT 2037
QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2038 ATTGATAGGAATGATCGATATGCCAACAGAACACCGTGNNNNNCGATATTAACTTGAGA 2097

```

```
QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
Db 2098 GGTCTTGTATGGCATCAAGGACAGTTTATGTGGGCACTGGCTGTGTTTCAACAGAAC 2157
QY 641 AlalIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
Db 2158 GCTCTATATGGTATGAGNNNNAAATTAAGCAAAAGAGGGTGGTTCTTGTGTCATCACTA 2217
QY 661 CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysSer 680
Db 2218 TGTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2274
QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
Db 2275 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2334
QY 701 GluGlyAlaGlyPheAspAspGlyLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db 2335 GAAAGCGCTGGATTTGATGATGAGAAATCACTTCTTATGTCTCAATGAGCTTGAGAG 2394
QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyValPro 740
Db 2395 AGATTGGCCAACTGCGAGCTTTGTGGTCCACTCTGATGGAATNNNNNNNNNNNNNNNN 2454
QY 741 GlnSerSerThrProGluSerLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db 2455 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2514
QY 761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrGlySerValThrGluAsp 780
Db 2515 GAGGACAAGATTGATGGGAACTGAGATTGGTGGATCTATGTTCTGTGACGGAAGAT 2574
QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
Db 2575 ATTCTACCTGGTTCAAGATGCACGACGAGCTGGCGGTGCATCTACTGCTGCTTAAG 2634
QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db 2635 CGGCGGCGCTTCAAGGATCGCTCCCATCAATCTCTCAGACCGTCTGACAGGTGCTC 2694
QY 821 ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrGly 840
Db 2695 CGGTGGGCTCTCGGTTCAGTGGAAATCTTTTCAGCGCGCATTTGNNNNNTATGGTACGG 2754
QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db 2755 TACGAGGACCGCTGAAGTTCTTGAGAGATTGCGCTACATCAACACACCATCTACCCG 2814
QY 861 LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2815 CTCACGTCCTCCCGCTCTCATTTACTGTATCTGCTGCTGCATCTGCTGCTCAGGGG 2874
QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
Db 2875 AAGTTCAATCAATCCAGAGATCAGCAACTTCGCTAGTATCTGGTTCACTCTCTTTCATC 2934
QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr 920
Db 2935 TCGATCTTCGCCACCGGTATCTCGAGATGAGGTGGAGCGCGGTGGGCATCGACGAGTGG 2994
QY 921 TrpArgAsnGluGlnPheTyrValIleGlyIleSerAlaHisLeuPheAlaValPhe 940
Db 2995 TGGAGGAACGAGCATGTTCTGGTTCATCGAGGCATCTCCGCCACCTCTTCGCCGTCTTC 3054
QY 941 GlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
Db 3055 CAGGCGCTCTCTCAAGTGTCTGCGGCATCGACCAACTTCACCGTCACTCCAGGCC 3114
QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeuIle 980
Db 3115 TCGGATGAAGACGCGCATCTTCGGGAGCTGTACATGTTCAAGTGGAGCAGACATCTCTGATC 3174
```

```
QY 981 ProProThrThrIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db 3175 CGCCACACCATCTGATCATCACTGGTGGGCTGTGTTGCGGCACTCTCTACGCC 3234
QY 1001 IleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPhePheAlaPheTyr 1020
Db 3235 ATCAACAGCGGTACCACTGCTGGGTGGCTCTTCGGCAAGCTCTTTCGCTTCTGG 3294
QY 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3295 GTGATCGTTCACTGTCACCGTTCCTCAAGGGTCTCATGGTGGCAGAACCGCACCCG 3354
QY 1041 ThrIleValIleValTyrAlaValLeuLeuAlaSerIlePheSerLeuLeuTyrValArg 1060
Db 3355 ACCATCGGTGTGTCTGGCGCATCTGCTGGCGTGCATCTCTCTGCTGTGGTTCGC 3414
QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3415 ATCGATCGTTCAACCAACCGGTCTACTGCCCGGATACTCGAACGTGTGGCATCACTGC 3474

RESULT 3
AY110079 3696 bp mRNA linear HTC 17-OCT-2002
LOCUS Zea mays CL1164_1 mRNA sequence.
DEFINITION Zea mays
ACCESSION AY110079
VERSION AY110079.1 GI:21214162
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3696)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 3696)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
location/Qualifiers
1..3696
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630055"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ALIGNMENT SCORES:
Pred. No.: 0 Length: 3696
Score: 4805.00 Matches: 895
Percent Similarity: 89.08% Conservative: 68
Best Local Similarity: 82.79% Mismatches: 112
Query Match: 83.16% Indels: 6
DB: 11 Gaps: 6
```

```
US-09-900-237A-30 (1-1080) x AY110079 (1-3696)
QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 254 ATGGACGGC--GGCGACGCCACGAATTCGGGGAAGCATGTGGCCGGCAGGTGTGCCAG 310
QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysaspVal 40
Db 311 ATCTGGCGACGGCGTGGGACCGCGCGGACGGCGACCTCTTACCOCCTCGCACGTC 370
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgGlyGlyThrGlnAla 60
Db 371 TCGCGCTTCCCCGTGTGCGGCCCATGCTACGAGTAGCGAGCGACGACGCCACCCAGGCG 430
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db 431 TGCCCCGAGTGAAGTACAAAGCGCCCAAAAGGGAGGCCACCAGTACACAGGTGAG 490
QY 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProHlaSerGlyThr 100
Db 491 GAAATATGAGATGTGAGTCTGACGATGTGAGTGACTACAACTACCAAGCATCTGGCAAC 550
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGlyGly 120
Db 551 CAGGATCAGAAGCAAAAGATTGCTGAGAGAACTGCTCCTTGGCGGACAAACTCACGTCGC 610
QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
Db 611 AGT---GATATTGGCCTGGCTAAGTATGACAGCGGTGAAATTGGCGATGGGAAGTATGAC 667
QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
Db 668 AGTGTGTGATCCCTCGTGATATATCCCGTCACTAACTCATAGCCAGATCTCAGGAGAG 727
QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
Db 728 ATTCCTGGAGCTTCCCTCGAT---CATATGATGTCTCTGTTGGGAACATTCGACGCGT 784
QY 181 Ala---ProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySer 199
Db 785 GGACATCAATTCCTTATGTAAATCATTCCTCAAACCCATCGAGGGAGTTCTCCGCTAGC 844
QY 200 IleGlyAsnValAlaIlePlysGluArgValAspGlyTyrLysMetLysGlnAspLysGly 219
Db 845 CTTGGCAATGTTCATCGAAAGAGGGTGGATGGATGGAAATGAAG---GATTAAGGT 901
QY 220 AlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaIleThrAsp 239
Db 902 GCNAATTCCTATGACCAATGGACACAGCATTTGCTCCATCAGAAGGGCGTGGAGTTGCTGAT 961
QY 240 IleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGln 259
Db 962 ATTTGATGCTTCTACTGATTATAACATGGAAGATGCCTTACTGAATGATGAAACTCGGCAA 1021
QY 260 ProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIle 279
Db 1022 CCTCTATCTAGAAAAGGCCAATTCCTTCATCCAGAAATAATCCGTACAGAAATGGTCATT 1081
QY 280 ValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArg 299
Db 1082 GTGCTAGCTTGGCTGTCTATGCATATCTTGGCTACCGTATCCACATCTCCTGTGAAC 1141
QY 300 AsnAlaTyrProIleThrLeuLeuSerValIleCysGluIleTyrPheAlaLeuSerTyr 319
Db 1142 AATGCATATCCACTGTGGCTTTTATCCGTATATGTGAGATCTGTGGTTGTCTGTCTGG 1201
QY 320 IleLeuAspGlnPheProLysTyrPheProIleAsnArgGluThrTyrLeuAspArgLeu 339
Db 1202 ATTTTGGATCAGTTCCCAAGTGGTCCCAATCAACCGTGAACATACATACCTTGATAGACTG 1261
QY 340 AlaLeuArgTyrAspArgGlyGluProSerGlnLeuAlaValAspIlePheVal 359
Db 1262 ATTTTGGATCAGTTCCCAAGTGGTCCCAATCAACCGTGAACATACATACCTTGATAGACTG 1321
1262 GCTTTAAGGTATGACCGAAGGTGAACCATCTCAATTAGTCTCTGTGTTGATANNNNNGTC 1321
360 SerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIle 379
1322 AGTACTGTGGATCCAATGAAGAGGCTCTCTGTGTCACTGCAAAATACGTGTCTTCCATC 1381
380 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer 399
1382 CTTGTCTGTGATTATCCGGTTGACAAGGTATCTTCTATGTTTCGATGATGAGGTGCT 1441
400 MetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTyrValProPhe 419
1442 ATGCTGACTTTTGTATCTCTCTGAAACTTCAGAGTTTGTAGAAAATGGGTTCGGTTC 1501
420 ValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAsp 439
1502 TGTAAAGAGTACAACATAGAGCTTANGGCCCGCGGAATGGTACTTTTGTCTCAGAAAATTGAT 1561
440 TyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGlu 459
1562 TACTTGAAGACAAAGTTCAAACCTCATTTGTGAAAGAACGCCGGNNNNNNNNNNNN 1621
460 TyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGlu 479
1622 NNNNNNNNNNCAAAAGTTCGTATCAATGCTGTGTAGCCAAAGCACNNNNNGTTCCCGAG 1681
480 GluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisPro 499
1682 GAGGATGATCATGCAAGATGGTACACCTTGGCCTGGGAAACAATACTAGGAGCATCCT 1741
500 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeu 519
1742 GGAATGATTTCAGTTTCTCGGTACAGTGGAGGGCTTGAAGTTCAAGGCAATGAACCT 1801
520 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAla 539
1802 CCTCGTTGGTTTATGTCTCGTGNNNNCGTCTCGGATTCGAAATCCAAATCACAAGAAGCT 1861
540 GlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeu 559
1862 GGTGCCATGAATGCATTTGTTGTGTATCAGTGTCTTACTAATGGCAATACATGTTG 1921
560 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPhe 579
1922 AATCTTGTATGTGACCACTACATCAATAATAGCAAGGCTCTTCGAGAAAGTATGTGTTCT 1981
580 LeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAsp 599
1982 CTTATGGACCCCAACCTAGGAAGGATGTCTGTATGTGCCAATTTCTCAGAGTTTGTAT 2041
600 GlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeu 619
2042 GGTATTTGATAGGAATGACCGATATGCAACAGAAACACTGTGNNNNCGATATTAACCTG 2101
620 ArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArg 639
2102 AGAGGCTTTCAGCGCATTCGAAGGCCAGTTTATGTGGAACTGGTTGTGTGTTTAAACAGA 2161
640 ThrAlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSer 659
2162 ACGGCTTATATGTTATGAGCTTCAGTCAAGNNNNNNNNNNNGCCAGGCTTCTTCTCTCG 2221
660 LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerAspLysLysLys 679
2222 CTTTGTNNNNNAAGNNNNNACGTCAAAATCTAAGAAG---AGTCGGAAGAAAGAAAG 2278
680 SerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGly 699
2279 TCACATAGACACGACAGACAGTTCTGTACCAAGTATTTAATTCGAAGATATAGAGGAAGG 2338
700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
2339 ATTGAAGGTTCTCAGTTTGTATGATGAGAAATCGCTGATTTATGTTCTCAATGAGCTTGGAG 2398
```

```

QY 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyVal 739
Db 2399 AAGAGATTGGCCAGTCCAGTGNNNNGTAGCCTCTACTCTGATGAATATGGTGTGTT 2458
QY 740 ProGlnSerSerThrProGluSerLeuLeuLeuGlyAlaLeHisValIleSerCysGly 759
Db 2459 CCACATCTGCAACTCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2518
QY 760 TyrGluAspLysSerGlyThrGluIleGlyTyrIleTyrGlySerValThrGlu 779
Db 2519 TATGAGCAAAACACTGACTGGGAACCTGAGATTGGGTGGATCTATGGTTCTGTACAGAA 2578
QY 780 AspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetPro 799
Db 2579 GACATTCTCACCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2638
QY 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
Db 2639 AAGCAGCAGCAGCTTTCAAGGGATCTGCTCTATCAACCTTCGGATCGTTGAATCAAGTG 2698
QY 820 LeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyr 839
Db 2699 CTTCGGTGGCTCTTGGTTCCATTGAAATCTTTTCAGCAGCATGTGCCATATGGTAT 2758
QY 840 GlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
Db 2759 GGCTATGGAGGCGCGCTTAAATCTCTGGAGAGATTGCTTATATCAACAACAATTTAT 2818
QY 860 ProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
Db 2819 CCACATCATCAATCCCGCTCTCTGTACTGATATGCCAGCATTTGTCTTCTCACT 2878
QY 880 GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPhe 899
Db 2879 GGGAAATTCATCATCCCAAGATTAGTAACCTTAGAGAGTGTGTGGTTTATATCGCTCTTT 2938
QY 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGlu 919
Db 2939 ATCTCAATCTTGGCAGTGTATCTCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2998
QY 920 TrpTyrArgAsnGluGlnPheTyrValIleGlyIleSerAlaHisLeuPheAlaVal 939
Db 2999 TGGTGGAGGAAACGACGACTCTGGGTCAATCTGGTATTTCTGCGCAITTTATTTGGCGTC 3058
QY 940 PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAnPheThrValThrSerLys 959
Db 3059 TTCAGGGTCTCCTGAAGGTGCTTGTGGTATCGACACGAGCTTCACCTGCACCTCTAAG 3118
QY 960 AlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTyrThrThrLeuLeu 979
Db 3119 GCCACTGACGAAGAAGGTGATTTTGGCAGCTCTACATGTTCAAGTGGACACGCTTCTG 3178
QY 980 IleProThrThrIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyr 999
Db 3179 ATCCCAACCAACCACTATTTTGTATCATCAACCTGCTGGCGTGTCTGCTGGCATTTCCTAC 3238
QY 1000 AlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3239 GCAATCAATAGCGGTACCATGTCATGGGACCTCTTTTCGGGAAGCTCTTCTTTTCGCTTC 3298
QY 1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
Db 3299 TGGGTGATTGTCACCTGTACCCCTTCTCAAGGGCCCTCATGGGAGACGAGAACCGCAGC 3358
QY 1040 ProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTyrVal 1059
Db 3359 CCGACCATTTGTGTGTCTGGGTCTATCTCTCTCGTCGATCTTTTCCCTGATGTGGGTT 3418
QY 1060 ArgValAspProPheThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
Db 3419 CGTATCGATTCATTCACACCGGGTCACTGGTCTCTGATATCGCGAAATGTGGCATCAAC 3478

```

```

QY 1080 Cys 1080
Db 3479 TGC 3481
RESULT 4
LOCUS AY110415
DEFINITION Zea mays CUL166_1 mRNA sequence.
ACCESSION AY110415
VERSION AY110415.1 GI:21214824
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3898)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 3898)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT these are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
source
1..3898
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630057"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3898
Score: 3902.00 Matches: 737
Percent Similarity: 79.27% Conservative: 131
Best Local Similarity: 67.31% Mismatches: 169
Query Match: 67.53% Indels: 58
DB: 11 Gaps: 22
US-09-900-237A-30 (1-1080) x AY110415 (1-3898)
QY 2 AspGlyAspAlaAspAla---LeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db 255 GACGGCCACCGCGCGGCCCGCGGTAGCCACCAAGAGATGCGAATGGCAGGTCTGCCAG 314
QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
Db 315 ATTTGGCGGACCTCTGTGGCGGTTTCAGCCACCTGGTGATGTCTTTGTTCCTGCAATGAG 374
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGlyThrGlnAla 60
Db 375 TGTGCTTCCTCTGTCGCCCTTGTCTAGTACGAGCGCAAGGAGGGAACCAATGCG 434
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80

```



```

QY 789 AlaAqGlyTyrArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAla 808
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2535 NNNAGGGGCTGCATCAATCACTACTGATGACACACACACCTTTTCAAGGGTTCTGCA 2594
QY 809 ProLeuAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGlu 828
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2595 CCAATCAATCTTCTGATGCTCTTAATCAGGTGCTCCGTTGGGCTTTGGGTCACTGGAA 2654
QY 829 IleLeuPheSerArgHisCysProLeuTyrTyrGlyTyrGlyArgLeuLysPheLeu 848
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2655 ATTCTGCTTAGCAGACATGCTCTATCTGGTATGTTTACCAATGGAGCATTAAGCTTTTG 2714
QY 849 GluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuVal 868
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2715 GAGAGCTGGCTTACATCAACACCATTTTATCCATCAACATCTGTTCCGCTTACGCC 2774
QY 869 TyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluLeuSer 888
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2775 TATTGTGTGCTCTCTGCTATCTCTTCTTACCAATAAATTTATCATTCCTGAGATTAGC 2834
QY 889 AsnLeuAlaSerIleTyrPheIleAlaLeuPheLeuSerIlePheAlaThrGlyLeu 908
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2835 AATTATGTGGGATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2894
QY 909 GluMetArgTyrSerGlyValGlyIleAspGluTyrTyrArgAsnGluLysPheTyrVal 928
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2895 GAGCTTAGATGAGTGTGTTGGCATTAAGATTTGGTGGAGAAATGACAGTTTGGGTT 2954
QY 929 IleGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAla 948
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
2955 ATTGTGTGACCTCTGCGCATCTCTTCTGAGTGTTCAGGGTCTGCTGAAAGTGTGGCT 3014
QY 949 GlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAla 968
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
3015 GGGATTGATACCAACTTACAGTTTACCTCAAAGGCATCTGATGAGGATGGGACTTTGCT 3074
QY 969 GluLeuTyrMetPheLysTyrThrThrLeuLeuIleProProThrThrIleLeuIle 988
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
3075 GAGCTATATGTTCAAGTGACACAGTTTGTCTCATCTCTCCGACACATGTTCTTGTCAAT 3134
QY 989 AsnMetValGlyValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyr 1008
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
3135 AACCTGTGCGAATGTTGGCAGGAATTTCTGATGCTATTAACAGTGGCTACCAATCTCTG 3194
QY 1009 GlyProLeuPheGlyLysLeuPhePheAlaPheTyrValIleValHisLeuTyrProPhe 1028
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
3195 GGTCCGCTCTTTGAAAGCTTCTTCTCTCGATCTGGGTGNTCTCCATCTGTACACGTTT 3254
QY 1029 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleValTyrAlaVal 1048
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
3255 CTCAGGGTCTCATGGGAGGACAGACCGCACCAACATTCGATCTGTGTCTTCATC 3314
QY 1049 LeuLeuAlaSerIlePheSerLeuLeuTyrValArgValAspProPhe-----Thr 1065
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
3315 CTTCTTCATCTATCTTCTCTCTGCTGTGGTGAAGATCGATCTTTCATCTCCCGGACA 3374
QY 1066 ThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
3375 CAGAAAGCTGTGCTCC-----TTGGGGCAATGTGGCGTCACTGC 3413

RESULT 5
AY104730
LOCUS AY104730 3783 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays P0100501 mRNA sequence.
ACCESSION AY104730
VERSION AY104730.1 GI:21207808
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.

```

```

REFERENCE 1 (bases 1 to 3783)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
AUTHORS Coe,E.H.
JOURNAL Direct Submission
TITLE 2 (bases 1 to 3783)
COMMENT Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
Location/Qualifiers
1..3783
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635765"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BAGs in conjunction with the Maize
Mapping Project"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3783
Score: 3817.00 Matches: 714
Percent Similarity: 77.92% Conservative: 140
Best Local Similarity: 65.15% Mismatches: 168
Query Match: 66.06% Indels: 74
DB: 11 Gaps: 16
US-09-900-237A-30 (1-1080) x AY104730 (1-3783)
QY 16 GlyAspValCysGlnIleCysAlaAspGlyThrThrLeuAspGlyAspValPhe 35
Db 291 GGCCAGGTGTGCCAGATTTCGGCGCAGCGCTGGGGCGCAACCCGACGGGGAGCGGTC 350
QY 36 ThrAlaCysAspValCysArgPheProValCysArgProCysTyrGluHisGluArgLys 55
Db 351 GTGGCCTGCACGAGTGGCGCTTCCCATCTGCGGGAGCTGTCTACGAGTACGAGCGCGC 410
QY 56 GluGlyThrGlnAlaCysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerPro 75
Db 411 GAGGGCAGCAGCAACTGCCCCAGTCGCAAGCCGCTTCAAGCGCTTCAAGGGTGGCG 470
QY 76 AlaIleArgGlyGluGluArgAspThrAspAlaAspAspGlySerAspPheAsnTyr 95
Db 471 CGCGTGGCCCGGACGAGGAGGAGCGGCTCGACGACCTGGAGAACGAGTTCACTGG 530
QY 96 ProAlaSerGlyThrGluAspGlnLysGlnIleAlaAspArgMetArgSerTyrArg 115
Db 531 -----AGCGACAGACGACGACTCCCATCTGCGCGAGTCCATGCTCCAGCGCCAC 581
QY 116 MetAsnThrGlyGlySerGlyAsnValGly-----HisProLysTyr 129
Db 582 ATGAGTACGCGCGCGCGCGCGAGCTCGACGGGTGCGCGAGCCATTCACCCCC----- 635
QY 130 AspSerGlyGluIleGlyLeuSerLysTyrAspSerGlyGluIleProArgTyrVal 149
Db 636 -----ATCCCCAAT-----GTT 647
QY 150 ProSerValThrAsnSerGlnMetSerGlyGluIleProGlyAlaSerProAspHisHis 169

```


Db	1017	TCAGCCGAATAATCCCTACAGGATGATTATCGTTATCCGGTTGGTGGTTGGTTTC	1076	Db	2097	ATTTATGTGGGTACTGGATGTGTTTTCAGACGCGCAGGCACCTGTATGTTTATGATGCTCT	2156
Qy	289	PheLeuHisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeuTrpLeuLeuSer	308	Qy	649	lleIysAlaLysLysPro-----GlyPheLeuAla	658
Db	1077	TTCTTCCACTACCGAGTGATGATCCGCGCAAGATGCAATTGGCATTTGGTGCATATCT	1136	Db	2157	---AAAACGAGAAGCCACCATCAAGAACTTGCAACTGCTGCCCAAGTGGTGCCTCTCT	2213
Qy	309	ValIleCysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPheProLysTrpPhe	328	Qy	659	SerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLys	678
Db	1137	GTAATCTGTGAAATCTGGTTTCGGATGCTCTGGATTCTTGATCGATCCCAAAGTGGT	1196	Db	2214	TCCTCTCGCAGCAGGAACAAGAACTACAAAACCAAGAGCGGAGGAAG	2273
Qy	329	ProIleAsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGlu	348	Qy	679	LysSer-----AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIle	696
Db	1197	CCATCGAGAGAGACTTACCTGACCGTTTGTCACTAAGTTTGACAAAGAGTCAA	1256	Db	2274	AAAAGATTATTTTCAAGAAAGCAAAAACCCATCTCTCGCATATGCTTTGGGTGAAATT	2333
Qy	349	ProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspProLeuLysGluPro	368	Qy	697	GluGluGluValGluGluValaclyPheAspAspGluLysSerValLeuMetSerGlnMet	716
Db	1257	CCCTCTCAGCTTGCTCCAATCACTTCTTTGTGATCGTTGATCCCAAGGAACCT	1316	Db	2334	GATGAAGGTGCTCCAGGTGCT-----GATATCGAAGAGCCCGGAATCGTAATCAACAG	2387
Qy	369	ProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLys	388	Qy	717	SerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyr	736
Db	1317	CCCTTGGTCACAGCAACACTGCTCTTCCATCTTCTGTGATTTATCCGGTTGAGAAG	1376	Db	2388	AAACTAGAGAAGAAATTTGGGCAGTCTTCTGTTTGTGCGCATCAACACTTCTTGAGAAC	2447
Qy	389	ValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGlu	408	Qy	737	GlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIle	756
Db	1377	GTCTCTGCTATGTTCTGATCATGCTGCTGCAATGCTTACGTTTGAACATTGTCTGAA	1436	Db	2448	GGAGGGACCTCGAAGAGCGCAAGTCCAGCTCTCTTCTGAAGGAAGCTATATGTTATC	2507
Qy	409	ThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAspIleGluProArg	428	Qy	757	SerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTrpIleTyrGlySer	776
Db	1437	ACATCTGAATTTGCAAGAAATGGGTCTCTTTTCAGCAAAAGATTTAAATATCGAGCTCGT	1496	Db	2508	AGCTCGGCTACGAAGACAAGACCAGCTGGGGAAGAGATTGGCTGGATTACCGATCG	2567
Qy	429	AlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGlnProSer	448	Qy	777	ValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyr	796
Db	1497	GCTCTGAGTGGTACTTCCACAGAGATAGACTACCTGAAAGACAGAGTTTGAAGTCTCA	1556	Db	2568	ATCACAGAGGATATCTTGACTGATTTAAGATGCATGCGCATGGCTGGCGGCTATTAC	2627
Qy	449	PheValLysAspArgAlaMetLysArgGluTyrGluGluPheLysIleArgIleAsn	468	Qy	797	CysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeu	816
Db	1557	TTTGTAGGAGAGAGGCGGATGAAGAGAGATACGAGGAATTCAGAGTAAGATCAAT	1616	Db	2628	TGCACTCCCGAAGCGGCTGCATTTCAAAGGTTCTGCGCTCTGAAACCTTTCCACCGCTCT	2687
Qy	469	AlaLeuValSerLysAlaLeuLysValProGluGluGlyTyrIleMetGlnAspGlyThr	488	Qy	817	AsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeu-PheSerArgHisCysPr	836
Db	1617	GCCTTGGTTGCAAAAGCCCAAAAGGTTCTCTGAGGAAGGATGACCAATGCAAGATGGAAGC	1676	Db	2688	CACCAAGTCTTCCGTGGGCCCTTGCGTCCGTCGAAATTTCTTTTCAGNAANCAATGGNC	2747
Qy	489	ProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHis	508	Qy	836	oleuTrpTyrGlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnTh	856
Db	1677	CCCTGCGCTGGAAAACAACGTACCGCATCTCTGGAATGATTTCAGTATTCTTGGCCAA	1736	Db	2748	ACTTTGGTACGATACGCGCGCGGCTAAAAATTCCTGGAAGGTTTCTTATATCACTC	2807
Qy	509	SerGlyGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGlu	528	Qy	856	rThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCy	876
Db	1737	AGTGGCGTCTGATGTGAAGGAATGAGTTGCTCGCTCGCTGTTTATGCTCGAGAGAA	1796	Db	2808	CATCGTTTATCCCTGACGCTCCATCTCTCTGGCTTACTGTACCTCGCTGCCATCTG	2867
Qy	529	LysArgProGlyPheGlnHisIleLysLysAlaGlyAlaMetAsnAlaLeuValArgVal	548	Qy	876	sLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIle	896
Db	1797	AAAGAGCCAGGTTATAACCATCACAGAAGGCTGGTGCCATGAATGCATGTGCTCGTGC	1856	Db	2868	CCTGCTCAGGGGAAGTTTATCACACAGAGCTTACCATGTGCGCAGTATCTGGTTTCAT	2927
Qy	549	SerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsn	568	Qy	896	eAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlu	916
Db	1857	TCTGCTGCTTATCAAAATCTGCTACATCCATTTGAACTTGACTGTGATCACTATCAAC	1916	Db	2928	GGCACTTTTTCATCTGCATCTCCGTGACCGCATCTCTGGAATTCAGGTGGAGTGGCGTGC	2987
Qy	569	AsnSerLysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyProGln	588	Qy	916	ylleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisIle	936
Db	1917	AATAGCAAGCCATAAAAGAGGCTATGTGTTTCATGATGATGCTCTTGTGGGGAAGAAA	1976	Db	2988	CATCGACGACTCGTGAGGAACGACGAGTCTCTGGGTTCATCGGAGGCGTTTCGGCGCATCT	3047
Qy	589	ValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArgTyrAla	608	Qy	936	uPheAlaValPheGlnGlyLeuLysValLeuAlaGlyIleAspThrAsnPheThrVa	956
Db	1977	GTGTGCTATGATACAGTTCCTCAGAGGTTTGATGGTATTGACAAATAATGATCATACGCT	2036	Db	3048	GTTCCGGGTTCACAGGGCTCTGTAAGGTGTTCGCGCGCATCGACACGAGCTTACCCT	3107
Qy	609	AsnArgAsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyPro	628	Qy	956	lThrSerLysAlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpTh	976
Db	2037	AAACGGAACGTTGCTCTTTTGTGATCAACATGAAGAAAGGTTTGACCGGTATTCAGGACCC	2096	Db	3108	GACGTCGAAGGCGGCGGACGACGAG---GAGTTCTCGAGCTGTACACGTTCAAGTGGAC	3164
Qy	629	ValTyrValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProPro	648	Qy	976	rThrLeuLeuIleProProThrThrIleLeuIleIleAsnMetValGlyValValAlaGlu	996
		...		Db	3165	CACCTTGTGATACCCCCGACACGCTCTCTCTGCTGAACTTCATCGGGGTGGTGGCCGG	3224

```

Qy 996 yThrSertYrAlaIleAanSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPh 1016
Db 3225 GATCTCGAACGGATCAACAACGGGTACAGTCTGGGGGCCCCCTGTTCGGAGACTCTT 3284
Qy 1016 ePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGl 1036
Db 3285 CTTGCGGTCTGGGTGATCGTCCACCTGTACCCGTTCCTCAAGGTCGTGGTGGGAGGCA 3344
Qy 1036 nAsnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLe 1056
Db 3345 GAACAGGAGCGCGAGTATCGTCTGCTGCTGCCATCTGCTGCTGCTGCTGCTGCTGCT 3404
Qy 1056 uLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCy 1076
Db 3405 CCTGTGGGTCTGGGTGACCGCTTCCTCGCCAGAGCGACGGCCGCTCCTGGAGGAGTG 3464
Qy 1076 sGlyIleAanCys 1080
Db 3465 TGGCCTGGACTGC 3477

RESULT 8
AY104236
LOCUS Zea mays PCO121439 mRNA sequence.
DEFINITION Zea mays PCO121439 mRNA sequence.
ACCESSION AY104236
VERSION AY104236.1 GI:21207314
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2872)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 2872)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizeMap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES
Location/Qualifiers
source
1..2872
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="maizeDB:637064"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Alignment Scores:
Pred. No.: 8,76e-309 Length: 2872
Score: 3525.00 Matches: 646
Percent Similarity: 88.42% Conservative: 87
Best Local Similarity: 77.93% Mismatches: 85
Query Match: 61.01% Indels: 11

```

```

DB: 11 Gaps: 6
US-09-900-237A-30 (1-1080) x AY104236 (1-2872)
Qy 258 ArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAanProTyrArgMet 277
Db 16 CGGCTACCTCTAAGTCGATAGTTCGATATCTCCAAAGAGCTTAACCTTATCGGATC 75
Qy 278 ValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnPro 297
Db 76 GTGATTGTTCCTCGGCTTATCATCTCTATGTTTCTTCTCAATATCGTATAACTCATCCA 135
Qy 298 ValArgAsnAlaTyrProLeuTrpLeuSerValIleCysGluIleTrpPheAlaLeu 317
Db 136 GTGGAAGATGCTTATGGGTTGTGGCTGTATCTCTTATTTGTGAAGTTTGGTTTGCCTTG 195
Qy 318 SerTrpIleLeuAspGlnPheProLysTrpPheProIleAanArgGluThrTyrLeuAsp 337
Db 196 TCTTGGCTTCTAGATCAGTTCCTCAAGTGGTATCTTATCAACCGTGAACCTTACTCGAT 255
Qy 338 ArgLeuAlaLeuArgTyrAspArgGluGluProSerGlnLeu-AlaAlaValAspIl 357
Db 256 AGACTTGCTTACATATGATAGGAGGGTGAGGCATCCCAAGTTGGGCTCCCAATCGATGT 315
Qy 357 ePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLe 377
Db 316 CTTTGTGTAGTACAGTGGATCCACTTAAGAACCTCTCTAATTACTGGCAACACTGTCTCT 375
Qy 377 uSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGl 397
Db 376 GTCCATTTCTGCTGGATTACCTCTGTGACAAAGTATCATGTTATGTTTCTGATGACGG 435
Qy 397 yAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpVa 417
Db 436 TTCAGTATGTTGACTTTTGAAGCGCTATCTGAAACCGCAGAGTTTCAAGAAATCGGT 495
Qy 417 lProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLy 437
Db 496 TCCCTTTTGAAGAAACACAAATATTGAACCTTAGGGCTCCAGAGTTTACTTTGCTCGAAA 555
Qy 437 sIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLy 457
Db 556 GATAGATTACTTAAGGRCANAATACAACTCTTTTGTGNAAGAAAGGCGGGCTATGAA 615
Qy 457 sArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysVa 477
Db 616 GAGGAGTGTGAAGAGTTCAAAGTACGATCGATGCCCTTGTTCAAAAGCGCAAAAAT 675
Qy 477 lProGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAs 497
Db 676 ACCTGAGGAGGGCTGGACCACTGGCTGATGGCACTCCCTGCGCTGGGAATAACCCCTAGA 735
Qy 497 pHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAs 517
Db 736 TCATCCMGGAATGATCCAGTATTCTTGGGCCACAGTGGTGGGCTTGACACGATGGGAA 795
Qy 517 nGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLy 537
Db 796 TGAGTTGCCACGGCTTCTTTATGTTTCTCGTGAAGAGGCGCAGGCTTCAGACCAACAA 855
Qy 537 sLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTy 557
Db 856 GAAGGCTGGTGGCCATGAATGCTTGTATTCGCTGATCAGCTGTCTGACGAAATGGTGTCTA 915
Qy 557 rMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMe 577
Db 916 TCTTCTTAATGTGGATTGTGATCACTACTTCAATAGCAGCAAAAGCTTTAGAGAGGCTAT 975
Qy 577 tCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnAr 597
Db 976 GTGTTTTCATGATGGATCCAGACTAGGAAGGAAAACTTGTCTATGTTTCCAGTTTCCACAAG 1035
Qy 597 gPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIl 617

```


Score: 2507.00 Matches: 462
 Percent Similarity: 84.44% Conservative: 70
 Best Local Similarity: 73.33% Mismatches: 84
 Query Match: 43.39% Indels: 14
 DB: 14 Gaps: 5

US-09-900-237A-30 (1-1080) x CD726831 (1-1874)

QY 320 IleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeu 339
 DB 2 GTGCTGGACCATGCTCCCTAAATGGTTACCTGTCGAGGACACATTTATTGACAGATTA 61
 QY 340 AlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheVal 359
 DB 62 TCAGCCAGGTTTGAAG 121
 QY 360 SerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIle 379
 DB 122 AGTACCGTTGATCCGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 181
 QY 380 LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer 399
 DB 182 CTGCTGGAGACTATCTCTGATAAAGTCTCTCTGATGTCGATGTCGATGTCGATGTCGAT 241
 QY 400 MetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaAlaArgLysTrpValProPhe 419
 DB 242 ATGCTCACATTGAATCTAGTTGAACAGCTGACTTTGCAAGGAATGGTTCATTC 301
 QY 420 ValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAsp 439
 DB 302 TGCAAAAAAATCTCCATTGAACCTCGAGCTCTGAGTTTACTTCTCCAGAAAAATAGAC 361
 QY 440 TyrLeuLysAspLysValGlnProSerPheValLysAspArgAtqAlaMetLysArgGlu 459
 DB 362 TATTTGAAGATTAAGTACAACCTCTCTTTGTAAGGAACGTAGAGCTATGAAGAGAGCT 421
 QY 460 TyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGlu 479
 DB 422 TATGAAGAGTTTAAAGTACGATCAATGCTTTGGTGGCAAGGCGCAAAACACCTGAT 481
 QY 480 GludlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisPro 499
 DB 482 GAAGGCTGGTCCATGCAAGATGGAACAGCTTTGGCCAGGAAACAATCCACGTCATCACCT 541
 QY 500 GlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeu 519
 DB 542 GGAATGATCAGGTTTTCTTGGAAATACCCGCGCCCATGACGTAGAGGAAATGAACCT 601
 QY 520 ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAla 539
 DB 602 CCTCGACTAGTCTACGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661
 QY 540 GlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeu 559
 DB 662 GGTCTGAAATAGCTCTGGTAAGAGTATCTGAGTCTCTCAAAATGCCCTTCATCTCT 721
 QY 560 AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPhe 579
 DB 722 AACCTTGACTGTGATCATGATTAACAATAGTCAAGCTATACGTAGGCAATGTGTTTC 781
 QY 580 LeuMetAspProAsnLeuLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAsp 599
 DB 782 TTGATGACCCACAAAGTTGGCCGAGATGATGCTATGTTCAAGTTTCTCAGAGGTTTGTAT 841
 QY 600 GlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeu 619
 DB 842 GGCATTGATCGAGTATCGATATATGCAATGCCAACCGGTTTCTTCGATGTAACATG 901
 QY 620 ArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArg 639
 DB 902 AAAGATTTGGATGTCATCCAGGACCAAGTTTATGTTGTCAGAGTTGTGTTTCAATAGA 961
 QY 640 ThrAlaIleTyrGlyTyrGluProProIleLys-----AlaLysLysProGlyPhe 656

DB 962 CAGGACACTTATGGCTACGAGACTCTCTCCATGCTAGCTTATCTAAGACATCCTCATCA 1021
 QY 657 LeuAlaSerLeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAsp 676
 DB 1022 TCTTGTCTCTGCTGTGGCTGTGCTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1081
 QY 677 Lys-----LysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu 693
 DB 1082 CCGACTGAGATTCAGAGAGATGCAAAAAAGAGAGAGCTTGATGCTGCAATCTTTAACCTC 1141
 QY 694 GluAspIleGluGluGlyValGluGlyAlaGlyPheAspAsp---GlyLysSerValLeu 712
 DB 1142 AGGGAATAGAT-----AATTATGATGAGTATGAGAGATCAATGCTG 1183
 QY 713 MetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThr 732
 DB 1184 ATTTCTCAACTGAGCTTTGAGAAAACTTCGGAATGCTGCTGCTGCTGCTGCTGCTGCTG 1243
 QY 733 LeuMetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLysGluAla 752
 DB 1244 CTAATGGAATAATGGCGAGTTGCGGAATCTGCAATCCCTCGACTTTGATCAAGGAAGCA 1303
 QY 753 IleHisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTrp 772
 DB 1304 ATTCATGTCATAGCTGTGGTTATGAAGAGAGTCCACTTGGGGAAAAAGAGATTGGTGG 1363
 QY 773 IleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrp 792
 DB 1364 ATATATGGCTCAGTGACTGAGGATATCTTAACGGGTTTCAAGATGTCATTTGCCAGGGTGG 1423
 QY 793 ArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeu 812
 DB 1424 AGGTCCATCTACTGTCATGCCATTTAGGCCAGCAITTCAAAGGGTCCGCAACCAATTAACCTT 1483
 QY 813 SerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSer 832
 DB 1484 TCTGATCGCTCCCAAGTTCTTCGATGGGACCTTGGATCTGTTGAGATTTTCTTAGC 1543
 QY 833 ArgHisCysProLeuTyrTrpTyrGlyTyr---GlyGlyArgLeuLysPheLeuGluArgPhe 851
 DB 1544 AGACACTGTCCATTTATGGATGGATTTTCAGCGCGCGCCCTCAAAATGGCTCCAAAGAATG 1603
 QY 852 AlaTyrIleAsnThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIle 871
 DB 1604 GCITTACATAAAACACCATTTGCTATCCCTTCACCTCGCTCCCTCTGTTGCTTACTGCTCA 1663
 QY 872 LeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAla 891
 DB 1664 TTGCTCGCATCTGCTGCTCACAGGAAGTTTCATCTTCCACGCTCTCGAACCTTAGCA 1723
 QY 892 SerIleTrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArg 911
 DB 1724 AGTACCTTTTCTCGGTCCTCTCTTGTCCATCATTTCTCAGAGTGTCTCGAGCTGGT 1783
 QY 912 TrpSerGlyValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGly 931
 DB 1784 TGGAGTGGTGTAGCATCGAAGATATATGCGGTAAACGAGCAATTTCTGGGTAATCGGAGC 1843
 QY 932 IleSerAlaHisLeuPheAlaValPheGln 941
 DB 1844 GTCCCGGCACATCTCTTTGCGCTCTCCAA 1873

RESULT 10
 CK272603
 LOCUS
 DEFINITION EST718681 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POD624 5' end, mRNA sequence.
 ACCESSION CK272603
 VERSION CK272603.1
 KEYWORDS EST,
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE AUTHORS TITLE JOURNAL COMMENT

1 (bases 1 to 916)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other ESTs: ESR718682
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: Arr TAG GTG ACA CTA TAG.

FEATURES

Location/Qualifiers
1..916
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="FOAD624"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="potato abiotic stress cDNA library"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed with withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.: 2,69e-132 Length: 916
Score: 1575.00 Matches: 283
Percent Similarity: 99.34% Conservative: 20
Best Local Similarity: 92.79% Mismatches: 0
Query Match: 27.26% Indels: 0
DB: 14 Gaps: 0

US-09-900-237A-30 (1-1080) x CK272603 (1-916)

QY 738 GlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaHisValIleSer 757
Db 1 GGTGTTCTCAATCGGTACCCGGAGACCTTTTGAAGAGCGTATTCATGATCAGT 60
QY 758 CysGlyTyrGluAspLysSerGluThrGlyThrGluIleGlyTrpIleTyrGlySerVal 777
Db 61 TGTGGTTATCAAGATAAATCAGATGGGAACTGAGATGGATGATCTATGTTCCGTC 120
QY 778 ThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCys 797
Db 121 ACAGAGGATATCTTCTTGGATTAAAGATGATCCCGTGGTGGCGATCTATTACTGT 180
QY 798 MetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsn 817

Db 181 ATGCCAAGAGAGACCCGCCCTTCAAAGGGTCAGCTCTATTAAATCTTTTCAGATCGTCTGAAC 240
QY 818 GlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeu 837
Db 241 CAAGTGCCTTCGATGGCGCTTTAGGTCAGTGGAAATCTTTTCAGTAGGCAATGTCTATA 300
QY 838 TrpTyrGlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThr 857
Db 301 TGGTATGATACATAATGAGCGGTTGAAGTGGTGGAGAGATTGCTTATGCAACCAACC 360
QY 858 IleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeu 877
Db 361 ATTTATCCAATCATTCCATCTCTATATATATGATGATGATGATGATGATGATGATGAT 420
QY 878 LeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAla 897
Db 421 CTATCTGGGAAATTCATTATCTCTAGATTAGTACCTTGTAGATCTGGTTTATATCC 480
QY 898 LeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIle 917
Db 481 CTCTTTCTTTCCATTTTCTGCTACTGCTATCTGGAGATGAGATGGAGTGGTGTGAAT 540
QY 918 AspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPhe 937
Db 541 GATGAATGCTGGAGAAATGAACAGTTTGGGTCATTGGTGGTGTGTCACCTCACCTGTT 600
QY 938 AlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThr 957
Db 601 GCGGCTTCCAAGGGTGTCTCAAGTGTGTGCTGTATGATACCAACTTACTGTGTCACA 660
QY 958 SerLysAlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThr 977
Db 661 TCCAAGGCATCAGATGAAGATGGGACTTTGGGAACTCTACTTGTTCAAATGGCAACT 720
QY 978 LeuLeuIleProThrThrIleLeuIleLeuIleAsnMetValGlyValValAlaGlyThr 997
Db 721 CTCTTTATACCCCCCACTACTCTCTCATTTGTAACCTGTAGAGTGTGGCCGGCAGTC 780
QY 998 SerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePhe 1017
Db 781 TCATATGCCATCAACAGTGGTTACCAATCATGGGGTCCCTCTTTGGTAAATATTCTTT 840
QY 1018 AlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsn 1037
Db 841 GCCTTTCTGGTGATGCTTCACCTTTACCTTTACCCCTTCCCAAAGGTCTCATGGGTCTG 900
QY 1038 ArgThrProThrIle 1042
Db 901 CGGACACCCCACTATC 915

RESULT 11

CF513822
LOCUS CF513822 924 bp mRNA linear EST 09-SEP-2003
DEFINITION CABUD0007_IIIF_B10 Vitis vinifera cv. cabernet sauvignon (Clone 8) Bud - CABUD Vitis vinifera cDNA clone Cabud0007_IIIF_B10 5', mRNA sequence.
ACCESSION CF513822
VERSION CF513822.1 GI:34545590
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis.
REFERENCE 1 (bases 1 to 924)
AUTHORS Goes da Silva, F., Iandolino, A., Lim, H., Baek, J., Jones, K. and Cook, D.
TITLE Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon' berries at various developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology

One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcook@cdavis.edu
 Seq primer: ACGGTACCGACATATGCC.

FEATURES

source

location/Qualifiers
 1..924
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon (Clone 8)"
 /db_xref="taxon:29760"
 /clone="Cabud0007_IIIF_B10"
 /sex="Hermaphrodite"
 /dev_stage="Pre-bloom (10-11 days before bloom)"
 /lab_host="DH5alpha"
 /clone_lib="Vitis vinifera cv. cabernet sauvignon (Clone 8) Bud- CABUD"
 /note="Organ: Bud; Vector: pDNR; Site: 1: SfiI; Site 2: SfiI; CABUD is a cDNA library of Vitis vinifera cv. 'Cabernet Sauvignon' Clone 8 dissected buds. Samples were collected May 13, 2002 from pre-bloom plants (10-11 days before bloom), pre-veraison. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
 5'-AAGCAGTGTGATCAACGACGAGTGGCCATTACGGCGGG-3' and
 5'-ATTCTAGAGCGGCGGCCACATG-dT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Alignment Scores:

Pred. No.: 2,7e-129 Length: 924
 Score: 1542.00 Matches: 277
 Percent Similarity: 98.37% Conservative: 25
 Best Local Similarity: 90.23% Mismatches: 4
 Query Match: 26.69% Indels: 1
 DB: 14 Gaps: 0

US-09-900-237a-30 (1-1080) x CFS13822 (1-924)

QY 775 GlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSer 794
 DB 2 GGGTCTGTACGGAAGATATCTTACAGATTCAAGATGTCATGCTCGGGTGGAGATCT 61
 QY 795 ValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAsp 814
 DB 62 ATTACTGCATGCCCAAGCGCGCCCTTCAAGGGGTCTGACCTATTAATCTTTTCAGAT 121
 QY 815 ArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHis 834
 DB 122 CGTCTGAACCAAGTCTTCGATGGCCCTCGGTCTGTGGAATCTTTTCAGTCGGCAC 181
 QY 835 CysProLeuTyrPyrGlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIle 854
 DB 182 TGTCTATCTCGTATGTTATGTTGGTGAAGGCTAAATGGCTTGAGATTCGCATATGTA 241
 QY 855 AsnThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuProAla 874
 DB 242 AACACCAACCATTTATCAATACCGCATCCCTCTCTGTGTTACTGCATCTCGCCGCT 301
 QY 875 IleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleTyr 894
 DB 302 GTTTGTTTGTCTACTGGAAGTTCATTTCCACAGATCAGTACATGTCAGATGATTGTTGG 361
 QY 895 PheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGly 914
 DB 362 TTTATATCCCTCTCTCCCTTCATCTTTGCTACCGGTATATTGGAGATGAGATGGAGTGGT 421
 QY 915 ValGlyIleAspGluTyrPyrArgAsnGluInPheTyrValIleGlyIleSerAla 934
 DB 422 GTTGGATCATGATGATGGTGAAGATGAGCAGTTTGGGTATTGGTGGCTGCTGCT 481

QY 935 HisLeuPheAlaValPheGlnGlyIleuLeuLysValLeuAlaGlyIleAspThrAsnPhe 954
 DB 482 CATCTTTTGTGTTTTCAGGCGCTGCTGAAAGTACTTGTGGAATTCATACCACTTC 541
 QY 955 ThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLys 974
 DB 542 ACTGTCACTTCAAGCTTCGGATGAAGACGGGGATTTTGGTGAATCTACATGTTTAA 601
 QY 975 TrrPthrLeuLeuIleProPthrThrIleLeuIleLeuIleAsnMetValGlyValVal 994
 DB 602 TGGACAACCTCTTCTCATCCACCCCAAGCTCCTCATATCACTCACTGGTGGTGTGT 661
 QY 995 AlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGlyLys 1014
 DB 662 GCAGGGATCTCTCATGCTATCAACAGTGGCTACCACTCATGGGGTCCACTCTTTGGCA 721
 QY 1015 LeuPhePheAlaPheTrrPthrValIleValHisLeuTyrProPheLeuLysGlyLeuMetGly 1034
 DB 722 CTCCTCTTGGCTTCTGGTGTATGCTTCACTCTATCCCTTCTCTCAAGGTCTGATGGA 781
 QY 1035 ArgGlnAsnArgThrProThrIleValIleValTrrPthrAlaValLeuLeuAlaSerIlePhe 1054
 DB 782 CGCCAGAACCGAACCACTTATGTTGTGTGTCGATTCCTCTGCTTCAATTTTC 841
 QY 1055 SerLeuLeuTrrPthrValArgValAspProPheThrThrArgLeuAlaGly-ProAsnIleGly 1074
 DB 842 TCGTGTGTTATGGTCAGAAATGATCCCTTCCACTAGATCAGTGGCCCCAGATGTTGA 901
 QY 1074 nThrCysGlyIleAsnCys 1080
 DB 902 GCAGTGTGGTATCAACTGC 920
 RESULT 12
 CK271294
 LOCUS
 DEFINITION
 EST717372 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POACY34 5' end, mRNA sequence.
 CK271294
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Solanum tuberosum (potato)
 Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
 Generation of ESTs from abiotic stressed potato tissue
 Unpublished (2003)
 Other ESTs: EST717373
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr. Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.
 FEATURES
 source
 1..977
 /location="Qualifiers
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="kennebec"
 /db_xref="taxon:4113"
 /clone="POACY34"
 /tissue_type="abiotic stress treated leaf and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="potato abiotic stress cDNA library"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: Solanum tuberosum var. Kennebec plants were
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation of
 the soil with 150 mM NaCl and tissues were harvested at

and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:
 Pred. No.: 8.9e-126 Length: 958
 Score: 1503.50 Matches: 281
 Percent Similarity: 95.00% Conservativeness: 23
 Best Local Similarity: 87.81% Mismatches: 12
 Query Match: 26.02% Indels: 4
 DB: 14 Gaps: 2

US-09-900-237A-30 (1-1080) x CK274934 (1-958)

QY 633 ThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleIysAla--- 651
 Db 1 ACTGGATGCTTCAATAGAACACCTTATATGTTATGACCTCCCAATTAAGCCAAAG 60
 QY 652 ---LysLysProGlyPheLeuAlaSerLeuCysGlyGly---LysLysLysAlaSerLys 669
 Db 61 CATAAGAGGCGAGGCTTCTCTCTCTGCTTCGCTGGATCAAGAAAGAGGTTCTTAAT 120
 QY 670 SerLysLysArgSerSerAspLysLysSerLysHisValAspSerValPro 689
 Db 121 TCAAGTAAAGAGGCTCAGACAAAGAAATCTAGTAAGAATGTTGATCCCACTGTGCCA 180
 QY 690 ValPheAsnLeuGluAspIleGluGluValGluGluValGluGluPheAspAspGluLys 709
 Db 181 ATATTCAATCTGGAGATATAGAGAGGAGTTGAAGTGTGCTGATTTGATGATGAGAAG 240
 QY 710 SerValLeuMetSerGlnMetSerLeuLulysArgPheGlyGlnSerAlaAlaPheVal 729
 Db 241 TCATCTCTCATGTCAAAATGAGCTGGAGAAGAGATTGGGCAATCGGCTGTTTGT 300
 QY 730 AlaSerThrLeuMetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeu 749
 Db 301 GCITTCACACTCATGGAGATGGTGGTGTCTCCTCAATCGGCTACCCAGAGACCCCTTTG 360
 QY 750 LysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGlu 769
 Db 361 AAGAGGCTATTCAATGTTATCAGTTGTTATGAGATAATCAGAAATGAGAACTGAG 420
 QY 770 IleGlyTrpIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAla 789
 Db 421 ATTGGATGGATCATGCTTCGTCACAGAGGATATTCTTATGATTTAAGATGATGCC 480
 QY 790 ArgGlyTrpArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAlaPro 809
 Db 481 CGTGGTGGCGATCTATATCTGATGATCCCAAGAGACCCGCTTCAAGGCTCAGCTCCT 540
 QY 810 IleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIle 829
 Db 541 ATTAATCTTTCAGATGCTGTAACCAAGTCTTCGATGGGCTTTAGGGTCAGTGGAAAT 600
 QY 830 LeuPheSerArgHisCysProLeuTyrGlyTyrGlyGlyArgLeuLysPheLeuGlu 849
 Db 601 CTTTTCAGTAGGCATTTCTCTATATGATGATGATACATGACGAGGTTGAAGTGGTGGAG 660
 QY 850 ArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuValTyr 869
 Db 661 AGATTGCTTATGTCACACCAACCATTTATCCCAATCACTTCCATTCCTTATATATAC 720
 QY 870 CysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsn 889
 Db 721 TGCATGCTCCAGCTATCTGCTACTTACTGGGAAATTCATTATCCCTCAGATTAGTAAC 780
 QY 890 LeuAlaSerIleTrpPheIleAlaLeuPheSerIlePheAlaThrGlyIleLeuGlu 909

Db 781 CTTCAGCATCTGGTTATATCCCTCTTCTTCCATTTCTACTGTTCTGGAG 840
 QY 910 MetArgTrpSerGlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIle 929
 Db 841 ATGAGATGAGTGGTGGTGGATTCATGATGATGGTGGAGAAATGACAGATTTGGGTCAAT 900
 QY 930 -GlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAla 948
 Db 901 GGGTGGTGTGTCAGCTCACCTGTTTGGCGTCTTCCAGGGTGTCTCAAGTGTCTTGT 958

RESULT 14

BG321258 978 bp mRNA linear EST 27-FEB-2001
 Zm04_03f11_R Zm04_AAPC_ECORC_cold_stressed_maize_seedlings Zea mays
 cDNA clone Zm04_03f11, mRNA sequence.

ACCESSION BG321258
 VERSION BG321258
 KEYWORDS EST
 SOURCE Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 978)
 Singh, J.A.; Wakui, K.; Couroux, P.; De Moors, A.; Harris, L.J.;
 Hattori, J.I.; Ouellet, T.; Robert, L.S.; Spott, D. and Tinker, N.A.
 Unpublished Sequence Tags from Cold-Stressed Maize Seedlings
 Contact: Singh, J.A.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-food Canada
 KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA
 O66, Canada
 Tel: (613) 759-1662
 Fax: (613) 759-1701
 Email: singhja@em.agr.ca.

FEATURES

source

1. 978
 Location/Qualifiers

/organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="CO328"
 /db_xref="taxon:4577"
 /clone="Zm04_03f11"
 /tissue_type="Leaf, crown"
 /note="Vector: Bluescript SK-/XhoI-EcoRI; Site 1: Eco RI;
 Site 2: Xho I; Lower temperature 50 C / hour from 22 to
 120C; bring to 50 in 1 hour from 120C. Leave at 50C 2
 days, photoperiod 16 hours. Light intensity was 125 uE-1.
 Library prepared by in vivo mass excision from amplified
 library."

ORIGIN

Alignment Scores:
 Pred. No.: 3.93e-122 Length: 978
 Score: 1463.50 Matches: 276
 Percent Similarity: 90.80% Conservativeness: 20
 Best Local Similarity: 84.66% Mismatches: 28
 Query Match: 25.33% Indels: 3
 DB: 12 Gaps: 1

US-09-900-237A-30 (1-1080) x BG321258 (1-978)

QY 559 LeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCys 578
 Db 3 TTGAATCTTGACCTGTTNACCACATACATATAGCAAGGCTCTTCGAGAAGCTATGTGC 62

QY 579 PheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPhe 598
 Db 63 TTCCTTATGGACCAACCTAGGAGGATGTCTGTBATGTCCAAATTTCTCAGAGTTT 122

QY 599 AspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsn 618

```

Db      123 GATGGTATTGATGAAATKACCGKATGCAACAGGACACATGTGTTTTCGATTATAAC 182
Qy      619 LeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsn 638
Db      183 TTGAGAGCKTTGATGCAATTCAGGGCCAGTTTATGTTKGAACATGGTTGTGTGTTTAAAC 242
Qy      639 ArgThrAlaIleTyrGlyTyrGluProProIleLeysAlaIleLeysLysProGlyPheLeuAla 658
Db      243 AGAAGCGGCTTTATATGTTTATKAGCTCCAGTCAAGAAAAAAGCCAGGCTTCTTCTCT 302
Qy      659 SerLeuCysGlyGlyLysLysAla-SerLysSerLysLysArgSerSerAspLysLys 678
Db      303 TCGCTTTTGGGGAGAGAAAGACGCGTCAAAATCTAAGAAG---AGCTCGAAAGAA 359
Qy      678 sLysSerAsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluCl 698
Db      360 GAAGTCAATAGACACAGCAGATCTCTGTACAGTATTATCTCGAAGATATAGAGGA 419
Qy      698 uGlyValGluGlyAlaGlyPheAspGluLysSerValLeuMetSerGlnMetSerLe 718
Db      420 AGGAATTTAAGGTTCTCAGTTTGTATGAGAAATCGCTGATTAATGTTCTCAATGAGCVT 479
Qy      718 uGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyCl 738
Db      480 KGAGAGAGATTTGGCCAGTCAGTCTTTTGTAGCCCTACTCTKATGGAATATGGCGG 539
Qy      738 vValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCy 758
Db      540 TGTTCACAAATCTGTACTCCGAGTCTCTTCTGAAGAAGCBATTCATGTCATCAGCTG 599
Qy      758 sGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValTh 778
Db      600 TKGCTATTAGAGCAAAACTGACTGGGGAAGTCTGAGATGGGTGGATCTATGGTCTGTAC 659
Qy      778 rGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMe 798
Db      660 AGAAGACATCTCACAGGATTCAGATGATGATGCBGAGGCTGGCGATCAATCTACTGCAT 719
Qy      798 tProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGl 818
Db      720 GCCTAARCSACCAGCTTTCAAGGGATCTGCTCTCYATCAACCTTTTCGGATCGTTGAATCA 779
Qy      818 nValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTr 838
Db      780 AGTGCYTCGGTGGCTCTTGGKTCATTKAAATCCCTTTTCAGCAGCATTTGCCATATG 839
Qy      838 pTyrGlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrI 858
Db      840 GTATGCTATTGAGCGCGCTTAAATTCCTGGAGARATTGGCTTTATCAACACAACAAAT 899
Qy      858 eTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLe 878
Db      900 TTATCCACTCACATCAATCCGCBCTCCYGTACTGCATATTGCCAGAT-TGTCTTCT 958
Qy      878 uThrGlyLysPheIle 883
Db      959 CACTGGGAAGTTCAATC 974

RESULT 15
CD439980      953 bp      mRNA      linear      EST 03-JUN-2003
LOCUS      EL01N0531D05.b EndospERM_5 Zea mays cDNA, mRNA sequence.
DEFINITION      CD439980
ACCESSION      CD439980.1 GI:31355623
VERSION      EST.
KEYWORDS      Zea mays
SOURCE      Zea mays
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE      1 (bases 1 to 953)
AUTHORS      Lai,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and

```

```

Messing,J.
Sequencing of the maize endospERM ESTs
Unpublished (2002)
Contact: Lai, Jinsheng
Dr. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: T3
Location/Qualifiers
1..953
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W22"
/db_xref="taxon:4577"
/tissue_type="EndospERM of 7-23DAP"
/clone_lib="EndospERM_5"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Alignment Scores:
Pred. No.: 4,19e-119 Length: 953
Score: 1430.00 Matches: 257
Percent Similarity: 91.64% Conservativeness: 28
Best Local Similarity: 82.64% Mismatches: 25
Query Match: 24.75% Indels: 1
DB: 14 Gaps: 0

US-09-900-237A-30 (1-1080) x CD439980 (1-953)
Qy 728 PheValAlaSerThrLeuMet-GluTyrGlyValProGlnSerSerThrProGluSe 747
Db 21 TTCATTCATCCACCTTTATGACACACCGTCGCATACCACCTTCAACAAACCCAGCTTC 80
Qy 747 rLeuLeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSerGluTyrGl 767
Db 81 TCTACTAAGGAAGCTATCCATGTCATCAGTTGTGGATATGAGGACAAACTGAATGGGG 140
Qy 767 yThrGluIleGlyTyrIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMe 787
Db 141 AAAAGAGATTTGGCTGATCTATGTTTCAAGGAGATATTCCTACCTGGGTTTAAAT 200
Qy 787 tHisAlaArgGlyTyrArgSerValTyrCysMetProLysArgProAlaPheLysGlySe 807
Db 201 GCATGCAAGGGCTGGCAATCAATCTACTGTCATGCCACCCAGACCTTGTTCAGGGTTC 260
Qy 807 rAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerVa 827
Db 261 TGCACCAATCAATCTTCCGATCGTCTTAATCAGGTGCTCCGTTGGGCTCTTCGGGTCAGT 320
Qy 827 lGluIleLeuPheSerArgHisCysProLeuTyrGlyTyrGlyArgLeuLysPh 847
Db 321 GGAATTTCTGTTAGTAGACATTTCTCTATCTGGTATGGTTTCAATGGAGCATTTGAGCT 380
Qy 847 eLeuGluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuLe 867
Db 381 TTTGGAGAGGCTGGCTTACATCAACACTATTGATATATCCAAATCACAATCCATCCGCTTAT 440
Qy 867 uValTyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluI 887
Db 441 TGCCTATTGTGTCTTCCGCTATCTGCTCTTACCAATAAATTTATCATCTCTCTAGAT 500
Qy 887 eSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeuSerIlePheAlaThrGlyI 907
Db 501 TAGCAATATTGCTGGGATGTTCTTCATCTCTTTTCGCCCTCCATTTTGCACCTGAT 560
Qy 907 eLeuGluMetArgTyrSerGlyValGlyIleAspGluTyrTyrArgAsnGluInPheTr 927
Db 561 ATTGGAGCTTAGATGGAGCGGTGTGGCATTTGAAGATTGGTGGAGAAATGAGCAGTTTG 620

```

```
QY 927 pValIleGlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValle 947
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
621 GGTATTGGTGACCTCTGCCCATCTCTCGCAGTGTCCAGGTCTGCTGAAAGTGT 680
QY 947 uAlaGlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGlyAspPh 967
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
681 GGCCTGGATGTATACCAACTTCACAGTTACCTCAAAGGCATCTGATGAGATGGGACTT 740
QY 967 eAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIleProProThrThrIleLeuIl 987
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
741 TGCTGAGCTATATGTCTCAAGTGGACCAAGTTTGCTCATTCCTCCGACCACCTGTTCTTGT 800
QY 987 eIleAsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSe 1007
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
801 CATTAACTGGTCGGAATGGTGGCAGGAATTTCGTATGCCATTACAGTGGCTACCAATC 860
QY 1007 rTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrPr 1027
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
861 CTGGGTCGGCTCTTTGGAAAGCTGTTCTTCGATCTGGTGATCCTCCTCACTCTCTACCC 920
QY 1027 oPheLeuLysGlyLeuMetGlyArgGlnAsn 1037
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
921 CCTTCTCAAGGTCATCGGAAGGCAAAAC 951
```

Search completed: August 22, 2004, 22:21:54
Job time : 6272 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 22, 2004, 13:37:22 ; Search time 935 Seconds
(without alignments)
4907.010 Million cell updates/sec

Title: US-09-900-237A-30

Perfect score: 5778

Sequence: 1 MDGADALKSRHGAGDVQC.....VDPFTTRIAGRNIQTGGINC 1080

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09900337/runat.18082004.081513.14627/app_query.fasta_1.1223
-DB=N_Geneseq_29Jan04 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09900337 @CGN 1.1.649 @runat.18082004.081513.14627 -NCPUS=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5434.5	94.1	3264	7	ADA69779	Ada69779 Rice gene
2	5423.5	93.9	3799	6	AAS16458	Aas16458 Corn cDNA
3	5421.5	93.8	3746	3	Aaz99512	Aaz99512 DNA encod
4	5421.5	93.8	3746	3	Aaz99527	Aaz99527 DNA encod
5	5421.5	93.8	3773	3	Aaz99494	Aaz99494 DNA encod
6	5169.5	89.5	3222	7	ADA69499	Ada69499 Rice gene
7	5143	89.0	3704	3	Aaz99533	Aaz99533 DNA encod
8	4530	78.4	3198	3	AAC49550	Aac49550 Arabidops

RESULT 1

ADA69779 ID ADA69779 standard; DNA, 3264 BP.

XX AC ADA69779;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 3102.

XX KW Plant; bacterial infection; fungal infection; viral infection; rice; gene; ds.

XX OS Oryza sativa.

XX PN WO2003000898-A1.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX DR WPI; 2003-175290/17.

XX PT Identifying at least one gene involved in plant resistance or response to pathogenic infection for conferring resistance or tolerance to a plant to

ALIGNMENTS

Aav06567 Arabidops
Aaz99500 DNA encod
Aaz99521 DNA encod
Aaz99506 DNA encod
Aaz99530 DNA encod
Aaz99515 DNA encod
Aaz99497 DNA encod
Aav06565 Arabidops
Aav06568 Arabidops
Aaz67114 Pinus rad
Aaz58265 Corn cell
Aaz99509 DNA encod
Aaz99524 DNA encod
Aaz99491 DNA encod
Aaz58263 Corn cell
Aaz58266 Corn cell
Aaz99518 DNA encod
Aaz99503 DNA encod
Aaz58268 Soybean c
Aav08373 Cellulose
Abz12754 Arabidops
Abz14725 Arabidops
Aav06566 Arabidops
Aac83798 Arabidops
Aas16455 Corn cDNA
Aav06563 Arabidops
Aac83247 Cellulose
Aav08372 Cellulose
Aav34432 Cotton ce
Aaa67145 Eucalyptu
Aac65448 Populus t
Aaz58270 Soybean c
Aav34433 Cotton ce
Aaz58271 Wheat cel
Aav08381 Cellulose
Aav06562 Arabidops
Aaz58269 Soybean c

9 4499 77.9 3614 2 AAV06567
10 3993.5 69.1 3725 3 AAZ99500
11 3993.5 69.1 3725 3 AAZ99521
12 3993.5 69.1 3725 3 AAZ99506
13 3981 68.9 3753 3 AAZ99530
14 3981 68.9 3753 3 AAZ99515
15 3981 68.9 3780 3 AAZ99497
16 3963.5 68.6 3603 2 AAV06565
17 3953.5 68.4 3673 2 AAV06568
18 3939 68.2 3851 3 AAA67114
19 3844.5 66.5 3786 3 AAZ58265
20 3844.5 66.5 3813 3 AAZ99509
21 3844.5 66.5 3813 3 AAZ99524
22 3820 66.1 3568 3 AAZ99491
23 3818 66.1 3776 3 AAZ58263
24 3818 66.1 3936 3 AAZ58266
25 3818 66.1 3969 3 AAZ99518
26 3818 66.1 3969 3 AAZ99503
27 3807.5 65.9 3517 3 AAZ58268
28 3777.5 65.4 3311 2 AAV08373
29 3666.5 63.5 3255 6 ABZ12754
30 3636 62.9 3255 6 ABZ14725
31 3633.5 62.9 3828 2 AAV06566
32 3583 62.0 3444 4 AAC83798
33 3535 61.2 2830 6 AAS16455
34 3518.5 60.9 8411 2 AAV06563
35 3514.5 60.8 7234 4 AAC83247
36 3482 60.3 3207 2 AAV08372
37 3477 60.2 3328 2 AAV34432
38 3461.5 59.9 3747 3 AAA67145
39 3337.5 57.8 3232 4 AAC65448
40 3234 56.0 2890 3 AAZ58270
41 2820 48.8 2306 2 AAV34433
42 2689 46.5 1734 3 AAZ58271
43 2674 46.3 2033 2 AAV08381
44 2629.5 45.5 2248 2 AAV06562
45 2627.5 45.5 2125 3 AAZ58269

PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.

XX Claim 6; SEQ ID NO 3102; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX Sequence 3264 BP; 784 A; 751 C; 852 G; 877 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3264
 Score: 5434.50 Matches: 998
 Percent Similarity: 96.60% Conservative: 52
 Best Local Similarity: 91.81% Mismatches: 30
 Query Match: 94.06% Indels: 7
 DB: Gaps: 2

US-09-900-237A-30 (1-1080) x ADA69779 (1-3264)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
 Db 1 ATGGACGGCGACGGCGATCGCGTGAAGTCGGGAGGACACGGGAGCGGGCGTCCAG 60
 QY 21 IleCysAlaAspGlyLeuGlyThrLeuAspGlyAspValPheThrAlaCysAspVal 40
 Db 61 ATCTGGCGGACGGCGTGGCGACGACGGCGGGCGCGCGTGTCTCGCGCGTGGACGTC 120
 QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
 Db 121 TGGCGCTTCCGGTGTGCGCCCTCTGCTACGAGTACGAGCGGCAAGGATGGCACCAGGCT 180
 QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
 Db 181 TGCCCCCAGTGCAGACCAAGTACAAGCGCCACAAGGGAGCGCGCGATCCCGTGGGAG 240
 QY 81 GluGlyAspAspThrAspAlaAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
 Db 241 GAAGCGGAGGACTACTGATGCTGAAGTGTGAGTACAGTACCACTACCCTGATCTGGCAGT 300
 QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
 Db 301 GCCGACCCAGAGCAGAGGATGCTGATAGGATCGGCAGTGGCGCATGAATGCTGGGGT 360
 QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
 Db 361 GGTGGAGACGTCGGCGGTCCCAAGTATGACAGTGGCGAGATCGGGCTCCACCAAGTATGAC 420
 QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
 Db 421 AGTGGCGAGATCCCTCGGGGATACATCCCTTCAGTCACTAATAGCCAGATCTCGGGAGAA 480
 QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
 Db 481 ATCCCTGGAGCTTCCCTGATCATCATATGATGCTCTCTACCGGAAACATTGGCAACGCT 540
 QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
 Db 541 GCTCCATTCCTTATGTGAACCATTCACCAATCCATCAAGGGAGTTCCTCGCAGCATT 600
 QY 201 GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyVala 220
 Db 601 GGAAATGTTCCCTGGAAAGAAGAGTTGATGGCTGGAACTGAAAGCAGGCAAGGGAGCA 660
 QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240

Db 661 ATTCCCATGACCAACGGGACAAAGCATTCGCCCTTCTGAAGGTCGGGGAGTTGGTGATATC 720
 QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
 Db 721 GATCATCCACTGATTAACAATATGAAGATGCTTACTGAAATGATGAATACTGCCAGCT 780
 QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
 Db 781 CTCTCTAGAAAGTTCCTTCCTTCAATCCAGAAATAATCCCTACAGAAATGTCATTGTT 840
 QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
 Db 841 CTGCGTGTGGTGTCTAAGCATTTTCTTCACTACCGTATTACGAATCCTTGGCAAT 900
 QY 301 AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle 320
 Db 901 CGGTATCCGCTCTGGCTTTTATCTGTATATGTAGATTGGTTGCTTGTCTTGATATA 960
 QY 321 LeuAspGlnPheProLysTrpPheProIleValThrAlaAsnThrValLeuSerIleLeu 340
 Db 961 TTGGATCAGTTCCTCCGAAGTGGTTTCCAAATCAACCGTGAAACCTACCTTGATAGCTGGCA 1020
 QY 341 LeuArgTyrAspArgGluGlyProSerGlnLeuAlaAlaValAspIlePheValSer 360
 Db 1021 TTGAGGTATGACAGAGAGGTGAGCCATCTCAGTTGGCTGCTGTGACATTTTGTGAGT 1080
 QY 361 ThrValAspProLeuLysGluProIleValThrAlaAsnThrValLeuSerIleLeu 380
 Db 1081 ACAGTCGACCCCTCAAGGAACCTCTCTGTGTACTGCGCAATACCGTGTGTCCTTCCT 1140
 QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
 Db 1141 GCTGTTGATTACCCAGTCGACAGGTCTCTGCTATGTATCTGACGATGGTGTGCAATG 1200
 QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
 Db 1201 CTGACTTTTGATGCAATTTGGCTGAGACTTCAGAGTTTGTAGAAAGTGGGTTCCTTCGTT 1260
 QY 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
 Db 1261 AAGAATATACCAATTGAGCCACAGAGCTCTGTGATGTACTTCTCCAGAAATTCGATTAC 1320
 QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr 460
 Db 1321 TTGAAGACAAAGTCCACCTTCATTTGTTAAAGACCGTCGTGCCATGAAGAGAAATAT 1380
 QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
 Db 1381 GAAGAAATCAAGATTAGGATAAATGGCTTGTGTCTAAGGCACAGAAAGTCCCGGAGGAG 1440
 QY 481 GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 500
 Db 1441 GGATGGATCATGCAAGATGGCACACCATGGCCAGGAAACAATATAGGGACCATCCTGGA 1500
 QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
 Db 1501 ATCATTTTCAGTTTTCCTTGGTTCACAGTGTGGCTTGTATACCGAGGGTATGAGCTTCCC 1560
 QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGly 540
 Db 1561 CGCTCGCTTACGTATCTGTGAGAAAGCGTCTGGGGTTTCAGCACCAAGAAAGCTGGT 1620
 QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
 Db 1621 GCCATGAATGCTCTTGTTCGTGTCTCAGCTGTCTTACCAATGGACAGTACATGTTCAAT 1680
 QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
 Db 1681 CTTGATTGTGATCACTACATCAACCAACAGCAAGGCTCTCGGGAGGCTATGTGCTTCCCT 1740
 QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600

```

Db 1741 ATGGATCCAAACCTAGGAAGAGTGTCTGTTATGTTTCAGTTCCTCCACAAAGGTTTCGATGTT 1800
QY 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
Db 1801 ATTGATAGGATGATCAATATGCGACAGAACACTGTGTTTTCATATTAACATTGAGG 1860
QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
Db 1861 GGTCTTGATGGCATCCAAAGACCAAGTTATGTGGGAACCTGTTGTATTCACAGAACT 1920
QY 641 AlaIleTyrGlyTyrGluProProlleLysAlaLysLysProGly---PheLeuAlaSer 659
Db 1921 GCTCTATATGTTATGAACCCCAATTAAGCAGAGAAAGAGGTTCTTTGTCATCA 1980
QY 660 LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLys 679
Db 1981 CTATGTGGGGCGGAAGAGGACAGCAAGTCAAAGAAAGAGGTTCGGACAGAAGAG 2040
QY 680 SerAsnLysHisValAspSerValProValPheAsnLeuGluAspIleGluGly 699
Db 2041 TCAAAACAAGCAGCTGGACAGTCTGTGCCAGTTTTCATCTTGAGATATAGAGGAGGT 2100
QY 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db 2101 GTTGAAGGTGCTGGATTCGATGATGAGAAATCACTTCTTATGCTCAAATGAGCTTGGAG 2160
QY 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyVal 739
Db 2161 AAGAGATTGGCCAGTCTGACGCGTTGTGCTCCACTCTCATGGAATATGTTGTTGTT 2220
QY 740 ProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGly 759
Db 2221 CCTCAATCTGCAACCCCAAGATCTCTTTGAAAGAGCTATCCATGATGATAGTTGGGC 2280
QY 760 TyrGluAspLysSerSerGluTyrGlyThrGlu-----IleGlyTyrIle 773
Db 2281 TATGAGGACAAGACGAATGGGGAACCTGAGGCTATGCAACTTGCAATGATTGGGTGGATC 2340
QY 774 TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArg 793
Db 2341 TACGGTTCGGTGACAGAGATATCTCACTGATTCAAGATGCAATGCGCGAGGTGGAGA 2400
QY 794 SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer 813
Db 2401 TCAATCTACTGCATGCCCAAGCCCGACCTTCAAGGGGCTGCTGCTCTATCAATCTTCA 2460
QY 814 AspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArg 833
Db 2461 GATCGTCTTAACCAAGTGCCTTCGGTGGGCACTGCTGTTGAAATCTTTTTCAGTCGC 2520
QY 834 HisCysProLeuTyrTyrGlyTyrGlyValArgLeuLysPheLeuGluArgPheAlaTyr 853
Db 2521 CATTTGCCATATGATCGGCTATGGAGGACGCTTAAGTTCTTGGAGAGATTTGCCCTAC 2580
QY 854 IleAsnThrThrIleTyrProLeuThrSerLeuLeuValTyrCysIleLeuPro 873
Db 2581 ATCAACACCACTATTTATTCATTGACATCGATCCCGCTTCTCATATACTGTGTTTGCCT 2640
QY 874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle 893
Db 2641 GCTATCTGTTTGTCTCACTGGGAAGTTCATCATCCAGAGATTAGCACTTTGCTAGTATT 2700
QY 894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSer 913
Db 2701 TGGTTCAATCTCTCTTCATTTCAATTTTGGCCACTGGTATCTCTTGAGATGAGTGGAGT 2760
QY 914 GlyValGlyIleAspGluTyrTyrArgAsnGluGlnPheTrpValIleGlyGlyIleSer 933
Db 2761 GGTGTTGGCATCGATGATGGTGGAGAAATGAACAGTTCTGGGTTATGGAGGTATATCT 2820
QY 934 AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn 953
Db 2821 GCGCATCTTTTGGCGTCTTCCAGGGTCTCTCAAGGTGCTTGTGATCGACACCAAT 2880
QY 954 PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe 973
Db 2881 TTCACTGTCACTCAAGGCTTCTGATGAAGATGGCACTTTTGTCTGAGCTCTACATGTTTC 2940
QY 974 LysTyrThrThrLeuLeuIleProProThrThrIleLeuIleIleAsnMetValGlyVal 993
Db 2941 AAGTGACAAACGCTTCTCATCCACGACGACATCTTGATCATTAACTGTCGTGTT 3000
QY 994 ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGly 1013
Db 3001 GTTGTGTTGATCTCATACGCGATCAACGCGCTACCATTCATGGGACCGCTCTTTTGC 3060
QY 1014 LysLeuPhePheAlaPheTyrValIleValHisLeuTyrProPheLeuLysGlyLeuMet 1033
Db 3061 AAGCTCTCTTTTGGCTTCTGGGTGATTGTCCACTTGTACCCCTTCTCAAGGGTCTTATG 3120
QY 1034 GlyArgGlnAsnArgThrProThrIleValIleValTyrAlaValLeuLeuAlaSerIle 1053
Db 3121 GGTGCGCAAAACCGCACCTCCGACCATCGTTGTTTGGGCAATCTTCTGCTTCGATC 3180
QY 1054 PheSerLeuLeuTyrValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle 1073
Db 3181 TTCTCATTTGCTGTGGTTTCGATCGATCCATTCACACCCGTTGTCCACCGGCCCATACC 3240
QY 1074 GlnThrCysGlyIleAsnCys 1080
Db 3241 CAACATGGCATCAACTGC 3261

```

RESULT 2
AAS16458
ID AAS16458 standard; cDNA; 3799 BP.
XX
AC AAS16458;
XX
DT 14-FEB-2002 (first entry)
XX
DE Corn cDNA encoding cellulose synthase Cqrae19/cesa-19.
XX
KW Corn; ss; cellulose synthase; Cdpqs45; cesa-3; Cqrae19; cesa-9;
KW stalk quality; improved stand; silage; pericarp; kernel hardening;
KW handling ability; transgenic plant.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 238..3799
FT /*tag= a
FT /product= "Cellulose synthase"
XX
XX WO200179516-A2.
XX
XX 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US011951.
XX
XX 14-APR-2000; 2000US-00550483.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Dhugga KS, Helentjaris TG;
XX
XX WPI; 2002-041338/05.
XX
XX P-PSDB; AAU10496.
XX
XX New cellulose synthase polypeptides and polynucleotides, useful in
XX improving stalk quality or silage, and in increasing concentration of
XX cellulose in the pericarp, hardening the kernel for improved handling
XX ability.
XX
XX Claim 4; Page 80-85; 88pp; English.
XX
XX The invention relates to isolated nucleic acids encoding two cellulose
CC

CC synthase proteins from corn, Cdp945 (cesa-3) and Cqrae19 (cesa-9). Also
 CC disclosed are a recombinant expression cassette comprising the
 CC polynucleotide (operably linked to a promoter) a host cell comprising the
 CC recombinant expression cassette and a transgenic plant comprising the
 CC recombinant expression cassette. The nucleic acid is useful in the
 CC improvement of stalk quality for improved stand or silage, and in the
 CC increased concentration of cellulose in the pericarp, hardening the
 CC kernel, and thus improving its handling ability. The nucleic acids may
 CC also be used as probes or amplification primers in the detection,
 CC quantification or isolation of gene transcripts, as probes in detecting
 CC deficiencies in the level of mRNA, for detecting gene mutations or
 CC allelic variants, for monitoring up regulation of expression or changes
 CC in enzyme activity in screening assays, for site directed mutagenesis,
 CC and in sense or antisense suppression of one or more genes in a host
 CC cell, tissue or plant. The polypeptides may be used in assays for enzyme
 CC agonists or antagonists, and as immunogen or antigen to obtain antibodies
 CC specifically immunoreactive with the protein. The present sequence
 CC encodes a corn cellulose synthase of the invention
 XX
 SQ

Sequence 3799 BP; 925 A; 931 C; 995 G; 944 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 0 Length: 3799
 Score: 5423.50 Matches: 996
 Percent Similarity: 96.48% Conservative: 46
 Best Local Similarity: 92.22% Mismatches: 37
 Query Match: 93.86% Indels: 1
 DB: 6 Gaps: 1

US-09-900-237A-30 (1-1080) x AAS16458 (1-3799)

QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
 Db 238 ATCGAGGGCGACCGCGAGTGGGAGTGGGAGGCGCGGGAGCGAGGTGTGCCAG 297
 QY 21 IleCysAlaAspGlyLeuGlyThrLeuAspGlyAspValPheThrAlaCysAspVal 40
 Db 298 ATCTGGCGGATGGCGTGGGCACCTACCGCGAGGAGACGCTCTCACCGCTCGACGTC 357
 QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
 Db 358 TGC GG GTTC CG GTT CCG CCG CCG TGT CAC GAG TAC GAG CGC AAG CGC AAG CGC 417
 QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
 Db 418 TGCCCCCAGTGC AAAA CAAGTACAAGCCCAAGGGGAGTCCAGCGATCCCGAGGGAG 477
 QY 81 GluGlyAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
 Db 478 GAAGGACGACATCTGATGCCGATGATGCTAGCGACTTCAACTACCTCGATCTGGCAAT 537
 QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
 Db 538 GACGACCAAGACAGAAATGTCTGACAGGTGCGCAGCTGGCCGATGAATGTCTGGGGGC 597
 QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
 Db 598 AGCGGGGATGTGGCCCGCCCAAGTATGACAGTGGTGAGATCGGGCTTACCAGTACGAC 657
 QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
 Db 658 AGTGGTGAGATCCCTCGGGGATACATCCCGTCAGTCACCTAACAGCAGATTCGGGAGAA 717
 QY 161 IleProGlyAlaSerProAspHisHisMetMetSerProThrGlyAsnIleSerArgArg 180
 Db 718 ATCCCTGTGTCTCCCTCGACCATCATATGATGCTCTACTGGGAAATTCGGCAGGGCGC 777
 QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
 Db 778 GCCCATTTTCCCTATATGAAATCATCAATCCGTGAGGGGAATTCCTGGTAGCGGT 837
 QY 201 GlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGlyAla 220

Db 838 GCGAATGTTGCTGGAAAGAGAGGGTGTGATGCTGGAAATGAACGACGACACGGAACA 897
 QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
 Db 898 ATTCCATGACGAATGGCACAGCATGTCTCCTCTGAGGGCGGGGTGTGGTGATAT 957
 QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
 Db 958 GATGCATCAACTGATTACCAACATGAAGATGCTTATTAAACGATGAAATCGCCAGCCT 1017
 QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
 Db 1018 CTATCTAGAAAGTTCCACTTCTCTCCAGATAAATCCATACAGAGTGTCTATGTG 1077
 QY 281 LeuArgLeuValIleuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
 Db 1078 CTACGATTGATTGTTCTAAGCATCTTCTTGCACTACCGGATCACAAATCTGTGCGTAAT 1137
 QY 301 AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle 320
 Db 1138 GCATACCCACTGTGGCTTCTATCTGTATATGTGAGATCTGGTTCCTCTTCTCTGATA 1197
 QY 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
 Db 1198 TTGGATCAGTTTCCAAAGTGGTTTCCATCAACCCGGAGACTTACCTTGATACACTGCA 1257
 QY 341 LeuArgTyrAspArgGluGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
 Db 1258 TTAAAGTATGACCGGAGAGGTGAGCCATCTCAGTGTGGTCTGTGACATTTTGTGAGT 1317
 QY 361 ThrValAspProLeuLysGluProIleValThrAlaAsnThrValLeuSerIleLeu 380
 Db 1318 ACTGTGCGACCAATGAAGAGCGCTCTCTGTGCTACCTGCCAATACCGTGTCTATCCATTC 1377
 QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMet 400
 Db 1378 GCTGTGAGCATCTCCTGTGGATAAGGTCTCTTGCTATGATCTGATGATGAGCTGCTATG 1437
 QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
 Db 1438 CTGACATTTGATGCACACTAGCTAGACTTCAGAGTTTGCTAGAAAATGGGTGCCATTTGTT 1497
 QY 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
 Db 1498 AAGAAGTACAACTTGAACCTAGAGCTCTCTGAATGGTACTTCTCCCAAGAAAATGATATAC 1557
 QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGluTyr 460
 Db 1558 TTGAAGGACAAAGTGCAACCTTCTATTGTTAAAGACCGCGGGCCATGAAGAGAAATAT 1617
 QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
 Db 1618 GAAGAATTCAAAATTAGGGTAAATGGCTTGTGCTAAGGCACAAAAAGTCCCTGAGGAA 1677
 QY 481 GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 500
 Db 1678 GGATGATCATGCAAGATGGCACCATGCCAGGAAACAATACAGGAGCCATCCTCTGGA 1737
 QY 501 MetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGluGlyAsnGluLeuPro 520
 Db 1738 ATGATTCAGGTTTTCTTGTGTCAGCTGGTGTCTTGATCTAGAGGGTAAATGAGTACCC 1797
 QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 540
 Db 1798 GGTGTTGTTATGTTTCTCGTGAAAAACGTCCTGGATTCAGCATCACAGAAAAGCTGGT 1857
 QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560
 Db 1858 GCCATGAATGCTCTGTCTCCGGCTCTCAGCTGTGCTTACCAATGACAATATCATGTTGAT 1917
 QY 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
 Db 1918 CTTGATGTGATCACTATACACACAGTAGGCTCTCAGGGAAGCTATGTGCTCTCTT 1977

Db 1995 CTTGATTGTGATCACTAATTAACAACAGTAGGCTCTCAGGAAGCTATGTGCTTCCT 2054
Qy 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
Db 2055 ATGACCCCTAACCTTAGGAAGAGTGTCTGCTACGTCCAGTTTCCCCAGAGATTGATGCG 2114
Qy 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
Db 2115 ATTGACAGGATGATCGATATGCCAACAGAACACCGTGTTTTTCGATATTAACTTGAGA 2174
Qy 621 GlyLeuAspGlyIleGlnClyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
Db 2175 GGTCTTGATGTCATCCAGGACCAAGTATATGCGAACTGCTGTGTTTTTCAACCGACA 2234
Qy 641 AlaIleTyrGlyTyrGluProProIleAlaIleAlaIleAlaIleAlaIleAlaIleSer 660
Db 2235 GCTCTATATGTTATGACGCCCCCAATTAAGCAGAGAGGTGGTTTTCTTCTCATCACTA 2294
Qy 661 CysGlyGlyIleLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLysSer 680
Db 2295 TGTGCGGTAGAGAGAGGCAAGCAATCAAGAGAG---GGCTCGACACAGAGATGCG 2351
Qy 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGluGlyVal 700
Db 2352 CAGAAGCATGTGGACAGTCTGTGCCAGTATTCACCTTGAAGATATAGAGAGGAGT 2411
Qy 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
Db 2412 GAAGGCGCTGGATTGTACGACGAGAAATCACTCTTATGTCTCAATGAGCGTGGAGAAG 2471
Qy 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro 740
Db 2472 AGATTGCGCAGTCCGACGCTTTGTGCTTCCTCTGATGGAGTATGGTGGTTCCT 2531
Qy 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
Db 2532 CAGTCCGCAACTCGGAGTCTCTCTGAAGAAGCTATCCATGTTATAAGCTGTGGCTAT 2591
Qy 761 GluAspLysSerGluTyrGlyThrGluIleGlyTrpIleTyrGlySerValThrGluAsp 780
Db 2592 GAGGACAAAGACTGAATCGGAACCTGAGATCGGTGGATCTACGGTCTCTGTGACAGAGAC 2651
Qy 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
Db 2652 ATTCTACCGGATTCAGATGCACGCGAGGCTGGCGGTGCTGATCTACTGATCCCAAG 2711
Qy 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
Db 2712 CGGCCAGCTTTCAGGGGTCTGCCCCCATCAATCTTTCGACCGCTCTGAACCGTCTC 2771
Qy 821 ArgTrpAlaLeuGlySerValGluIleuPheSerArgHisCysProLeuTyrTyrGly 840
Db 2772 CGGTGGGCTCTTGGGTCCGTGGAGATCCTCTTCAGCCGGCAGTCCGCCCTGTGTGACGGC 2831
Qy 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
Db 2832 TAGGGGGCGGCTCAGTTCTTGGAGAGATTCGGTACATCAACACCACTACCCG 2891
Qy 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
Db 2892 CTCAGTCCATCCGCTTCTCATCTACTGATCTCTGCCCGCATCTCTGTCTGCCCGGA 2951
Qy 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTrpPheIleAlaLeuPheLeu 900
Db 2952 AAGTTTCATATTCAGAGATCAGCAACTTCGCCAGCATCTGGTTTCATCTCCCTCTTCA 3011
Qy 901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrp 920
Db 3012 TCGATCTTCGCCACGGCATCTTGGAGATGAGGTGGAGCGGGTGGCATCGACGAGTGG 3071
Qy 921 TrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaValPhe 940

Db 3072 TGGAGAACAGCAGCAGTTCTGGGTGATCGGGGCGATCTCCGGCACCTCTTCGCCGTGTT 3131
Qy 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerIysAla 960
Db 3132 CAGGCGCTGTCTCAAGGTGCTGGCGGCGATCGACCAACTTCACGTCACCTCCAAGGCC 3191
Qy 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
Db 3192 TCGACAGAGGACGGCGACTTCGCGAGCTGTACATGTTCAAGTGGACGACGCTCCTGATC 3251
Qy 981 ProProThrThrIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
Db 3252 CGGCCACCACTCCTGATCATCACTGTGCGGCTGCTGCGCGCATCTCTACGCC 3311
Qy 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
Db 3312 ATCAACAGCGGATACCACTGCGGGCGCGCTCTTCGCAAGCTCTTCTCGCCTTCGG 3371
Qy 1021 VallileValHistuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
Db 3372 GTCATGTCACCTGTATCCCGTTCCTCAAGGCGCTCATGGGCGAGCAGAACCGCACCCG 3431
Qy 1041 ThrIleValIleValTrpAlaValLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
Db 3432 ACCATGTCGTGCTGCGTGGCCATCTGCTGGCGTCCATCTTCTCTGCTGTGGTTCGC 3491
Qy 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
Db 3492 ATCGACCCCTTCCACCGCGGTCACTGGCGCGGATACCCAGACGCTGTGGCATCAACTGC 3551
RESULT 4
AAZ99527 standard; DNA; 3746 BP.
XX
AC AAZ99527;
XX
DT 03-JUL-2000 (first entry)
XX
DB DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 321..3449
FT /*tag= a
FT /product= "cellulose synthase"
FT /transl_except= (pos: 1800..1802, aa: Xaa)
FT /note= "no termination codon given; Xaa is an unspecified amino acid"
XX
FN W0200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US018760.
XX
PR 17-AUG-1998; 98US-0096822P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX WPI; 2000-224343/19.
DR P-PSDB; AAY84119.
XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants and
PT to produce transgenic plants expressing the novel protein.
XX
PS Claim 1; Page 176-181; 119pp; English.

The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein

Alignment Scores:			
Pred. No.:	0	Length:	3746
Score:	5421.50	Matches:	998
Percent Similarity:	96.67%	Conservative:	46
Best Local Similarity:	92.41%	Mismatches:	33
Query Match:	93.81%	Indels:	3
DB:	3	Gaps:	2

US-09-900-237A-30 (1-1080) x AAZ99527 (1-3746)

QY	1	Met	Asp	Gly	Val	Asp	Ala	Asp	Ala	Leu	Leu	Ser	Gly	Arg	His	Gly	Val	Asp	Val	Cys	Gln	20
DB	321	AT	G	A	G	G	G	C	G	C	G	G	A	T	G	G	A	G	G	C	G	380
QY	21	Ile	Cys	Ala	Asp	Gly	Leu	Gly	Thr	Thr	Ileu	Asp	Gly	Asp	Val	Phe	Thr	Ala	Cys	Asp	Val	40
DB	381	AT	C	T	G	C	G	C	G	A	C	C	A	C	G	G	G	A	G	G	G	440
QY	41	Cys	Arg	Phe	Pro	Val	Cys	Arg	Pro	Cys	Tyr	Glu	His	Glu	Arg	Gly	Ser	Glu	Gly	Thr	Gln	60
DB	441	T	G	C	G	G	T	T	C	C	G	G	T	G	T	C	A	C	A	G	A	500
QY	61	Cys	Leu	Gln	Cys	Leu	Thr	Ile	Ser	Tyr	Leu	Arg	His	Arg	Gly	Ser	Pro	Ala	Ile	Arg	Gly	80
DB	501	T	G	C	C	C	C	A	G	T	C	A	A	G	C	C	A	A	G	G	G	560
QY	81	Glu	Gly	Asp	Asp	Thr	Asp	Ala	Asp	Asp	Gly	Ser	Asp	Phe	Asn	Tyr	Pro	Ala	Ser	Gly	Thr	100
DB	561	G	A	A	G	G	A	G	A	C	A	C	A	T	G	A	T	G	A	T	G	614
QY	101	Glu	Asp	Gln	Leu	Gln	Ile	Ser	Ile	Ala	Asp	Arg	Met	Arg	Ser	Tyr	Arg	Met	Asn	Thr	Gly	120
DB	615	G	A	G	A	C	C	A	A	G	A	T	T	C	C	G	A	A	T	G	C	674
QY	121	Ser	Gly	Asn	Val	Gly	His	Pro	Leu	Ser	Tyr	Asp	Ser	Gly	Glu	Ile	Gly	Leu	Ser	Tyr	Asp	140
DB	675	A	C	G	G	G	G	A	T	T	G	T	G	T	G	C	C	C	A	A	G	734
QY	141	Ser	Gly	Glu	Ile	Pro	Arg	Gly	Tyr	Val	Pro	Ser	Val	Thr	Asn	Ser	Gln	Met	Ser	Gly	Glu	160
DB	735	A	G	T	G	G	C	A	G	A	T	T	C	A	T	G	A	C	A	T	G	794
QY	161	Ile	Pro	Gly	Val	Ser	Pro	Asp	His	His	Met	Ser	Pro	Thr	Gly	Asn	Ile	Ser	Arg	Arg	180	
DB	795	A	T	C	C	C	T	G	G	T	T	C	C	C	A	T	A	T	A	T	G	854
QY	181	Ala	Pro	Phe	Pro	Tyr	Val	Asn	His	Ser	Pro	Asn	Pro	Ser	Arg	Glu	Phe	Ser	Gly	Ser	Ile	200

QY 561 LeuAspCysAspHisTyrIleAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
 Db 1995 CTTGATTGTGATCACTACATTAAACAAGTACCTCAGGAAGCTATGTGCTTCCTT 2054
 QY 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
 Db 2055 ATGGACCCCTAACTAGGAAGAGTGTCTGCTACGTCCAGTTTCCCAGAGATTGATGGC 2114
 QY 601 IleAspArgAsnAspArgTyrIleAlaAsnArgAsnThrValPhePheAsnIleAsnLeuArg 620
 Db 2115 ATTGACAGAAATGATCGATATGCCAACAGAAACACCGTGTGTTTTCGATATTAACTTGAGA 2174
 QY 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
 Db 2175 GGTCTTGATGGATCCAGGACAGTTTATGTGCGAAGTGGTGTGTTTCAACCGAACA 2234
 QY 641 AlaIleTyrGlyTyrGlnProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
 Db 2235 GCTCTATATGTTATGAGCCCCCAATTAAAGCAGAAAGAGGTGGTTTCTTGTCACTACTA 2294
 QY 661 CysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysSer 680
 Db 2295 TGTGGCGGTAGGAAGAGCAAGCAAGCAAAATCAAAGAAG---GGCTCGGACAAGAAGAAGTCG 2351
 QY 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGlnGlyVal 700
 Db 2352 CAGAAGCATGTGGACAGTTCTGTGCGAGTATTCACCTTGAGATATAGAGGAGGAGTT 2411
 QY 701 GluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGluLys 720
 Db 2412 GAAGCGCTGGATTTCAGCAGCAGAAATCACTTCTTATGTCTCAATGAGCTGGAGAAG 2471
 QY 721 ArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro 740
 Db 2472 AGATTGCGCCAGTCGCGAGCGTTGTTCCTCCACTCTGATGGAGTATGGTGGTTCCT 2531
 QY 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
 Db 2532 CAGTCCGCACTCCGGAGTCTCTCTGAAAGAGCTATCCATGTTATAGCTGTGGCTAT 2591
 QY 761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
 Db 2592 GAGGACAAGACTGAATGGGAACCTGAGATCGGGTGGATCTACGGTTCGTGTGACAGAGAC 2651
 QY 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
 Db 2652 ATTCTACCGGATTCAGATGACGCGAGGCTGGCGGTGCGATCTACTGATGCCCAAG 2711
 QY 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
 Db 2712 CGGCCAGCTTTCAGGGGTCTGCCCCCACTCAATCTTTCGGACCGTCTGAACAGGTGCTC 2771
 QY 821 ArgTrpAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTrpGly 840
 Db 2772 CGGTGGGCTCTTGGGTCGTGGAGATCTCTTTCAGCGGCACTGCCCTCTGTGTACGGC 2831
 QY 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
 Db 2832 TACGAGGGCGGCTCAAGATTCTCTGGAGAGATTCGCGTACATCAACACCCATCTACCCG 2891
 QY 861 LeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
 Db 2892 CTCAGTCCATCCCGCTTCTCATCTACTGCATCTGCGCGCATCTGTCTGCTCACCGGA 2951
 QY 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
 Db 2952 AAGTTCATCTCCAGAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTCTTCATC 3011
 QY 901 SerIlePheAlaThrGlyIleLeuGluMetArgTrpSerGlyValGlyIleAspGluTrp 920
 Db 3012 TCGATCTTCGCGACGGCATCTCGAGATGAGGTGGAGCGGGTGGCGATCGACGAGTGG 3071

QY 921 TrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisIleuPheAlaValPhe 940
 Db 3072 TGGAGGAACGAGCAGCTTTCGGGTGATCGGGGCGATCTCCGCGCACCTCTTCGCGGTGTC 3131
 QY 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
 Db 3132 CAGGGCTGCTCAAGGTGTGCGCGGATCGACCAACTTCACCGTCACTCCCAAGGCC 3191
 QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
 Db 3192 TCGGACGAGGACGGGACTTCGCGGAGCTGTACATGTTCAAGTGGACGAGCTCCTGATC 3251
 QY 981 ProProThrThrIleIleIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
 Db 3252 CGCCCCACACCATCTGATCATCAACCTGGTGGCGCTGTCGCGGCACTCTCTACGCGC 3311
 QY 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
 Db 3312 ATCAACAGGGATACCACTGCTGGGCGCGCTCTTCGGCAAGCTCTTTCGCCCTTCGG 3371
 QY 1021 VallleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
 Db 3372 GTCATGCTCCACTGTACCCGTTCTCAAGGCGCTCATGGGCGAGCAGAACCGCACCGC 3431
 QY 1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
 Db 3432 ACCATGCTGCTGCTGCGGCATCTGCTGGCGCTCCATCTTCTGCTGTGGGTTCGC 3491
 QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
 Db 3492 ATCGACCCCTTCACACCGCGTCACTGGCGCGGATACCCAGACGTGGCATCAACTGC 3551

RESULT 5
 AAZ99494
 ID AAZ99494 standard; DNA; 3773 BP.
 XX
 AC AAZ99494;
 XX
 DT 03-JUL-2000 (first entry)
 XX
 DE DNA encoding a maize cellulose synthase.
 XX
 KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
 KW transgenic plant; plant breeding marker; ss.
 XX
 OS Zea mays.
 FH Key Location/Qualifiers
 CDS 337..3565
 FT /*tag= a
 FT /product= "cellulose synthase"
 FT /transl_except= (pos: 1817..1819, aa: Xaa)
 FT /note= "no termination codon given; Xaa is an unspecified
 FT amino acid"
 XX
 PN WO200009706-A2.
 XX
 PD 24-FEB-2000.
 XX
 XX 16-AUG-1999; 99WO-US018760.
 XX
 XX 17-AUG-1998; 98US-0096822P.
 XX
 XX (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
 XX
 DR WPI; 2000-224343/19.
 DR P-PSDB; AAY84108.
 XX
 PT New genes which encode maize cellulose synthase polypeptides in plants
 PT useful for modulating the expression of cellulose synthase in plants and
 PT to produce transgenic plants expressing the novel protein.

XX
PS
XX
Claim 1; Page 90-94; 119pp; English.
The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting or mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein
SQ Sequence 3773 BP; 863 A; 988 C; 1017 G; 904 T; 0 U; 1 Other;
Alignment Scores:
Pred. No.: 0 Length: 3773
Score: 5421.50 Matches: 998
Percent Similarity: 96.67% Conservatives: 46
Best Local Similarity: 92.41% Mismatches: 33
Query Match: 93.83% Indels: 3
DB: 3 Gaps: 2
US-09-900-237A-30 (1-1080) x AA299494 (1-3773)
QY 1 MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
DB 338 ATGGAGGGCGACGCGGACCGCGTGAAGTCGGGAGCGCGGTGGCGGACAGGTGTGCCAG 397
QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
DB 398 ATCTGCGGCGACGGCGTGGGCACCAACGCGGAGGGGACGCTCTTCGCGCGCTCGCAGCTC 457
QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
DB 458 TGGCGGTTTTCGGGTGTGCGGCCCTGCTACGAGTACGAGCGCAAGGACGCGACGCGGCG 517
QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
DB 518 TGCCCCAGTGGCAAGACCAAGTACAGCGCCACAGGGGAGCCCGCGGATCCGTGGGGAG 577
QY 81 GluGlyAspAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr 100
DB 578 GAAGGAGACGACACTGATGCGGAT-----AGCGACTTCAATTACCTTGCATCTGGCAAT 631
QY 101 GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly 120
DB 632 GAGGACGAGAAGCAGAAGATTCGCGACAGATCGCGAGCTGGCGCATGAACGTTGGGGGC 691
QY 121 SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp 140
DB 692 AGCGGGAGTGTGGTCCGCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAAGTATGAC 751
QY 141 SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu 160
DB 752 AGTGGCGAGATTCTCGGGGATACCCATCAGTCACCTAACACCCAGATCTCAGGAGAA 811
QY 161 IleProGlyAlaSerProAspHisHisMetSerProThrGlyAsnIleSerArgArg 180
DB 812 ATCCCTGGTGTCTCCCTCGACCATCATATGATGTGCCCACTGGGAAACATTGGCAAGCGT 871

QY 181 AlaProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIle 200
DB 872 GCTCCATTTCCTATGTGAACCATTCGCAAAATCGGTCAAGGGAGTCTCTCGTAGCATT 931
QY 201 GlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAla 220
DB 932 GGGAAATGTGCTGGAAGAGAGGGTGTGATGGCTGGAATAATGAGCAGCAAGGGGAGC 991
QY 221 IleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIle 240
DB 992 ATTCCTCATGCAAGATGGCAGACAGCATTGTCTCCTGTAGGGTGGGGTGTGGTGATAT 1051
QY 241 AspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnPro 260
DB 1052 GATGCATCAACTGATTACACATGGAAGATGCTTATTGAACGACGAAATCTGCAGCGCT 1111
QY 261 LeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIleVal 280
DB 1112 CTATCTAGAAAGTTCCTCCTCCAGATAAATCCATACAGGATGGTCATGTG 1171
QY 281 LeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsn 300
DB 1172 CTGCGATGATGTCTTAAGCATCTCTTGCACTACCGTATCACAATCTGTGCGCAAT 1231
QY 301 AlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIle 320
DB 1232 GCATACCCATTATGGCTTCTATCTGTATATGTGAGATCTGGTTGCTCTCTTCGTGGATA 1291
QY 321 LeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAla 340
DB 1292 TTGGATCAGTTCCTCAAGTGGTTCATCAACCGGGAGACGTACCTTGTATAGGCTGGCA 1351
QY 341 LeuArgTyrAspArgGlyGluProSerGlnLeuAlaAlaValAspIlePheValSer 360
DB 1352 TTAAGTATGACCGGAAAGGTGAGCCATCTCAGTTCGCTGCTGTGACATTTTCGTGAGT 1411
QY 361 ThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeu 380
DB 1412 ACAGTCGACCCCAATGAAGGAGCGCTCTCTGTGTCACATGCGCAATACCGTGTATCTCTT 1471
QY 381 AlaValAspTyrProValAspLysValSerCysTyrValSerAspGlyAlaSerMet 400
DB 1472 GCTGTGGATACCTGTGGATAAGGTCTCTGTCTATGTATCTGATGATGAGGTGGATG 1531
QY 401 LeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheVal 420
DB 1532 CTGACATTCATGACTAGCTAGCTGAGACTTCAGAGTTCAGAAAATGGGTACCATTTGTT 1591
QY 421 LysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyr 440
DB 1592 AAGAAGTACAACATTTGAACCTAGAGCTCTGAAATGGTACTTCTCCAGAAAATGTATTAC 1651
QY 441 LeuLysAspLysValGlnProSerPheValLysAspArgArgAlaMetLysArgGluTyr 460
DB 1652 TTGAGGACAAAGTGCACCCCTTCATTTGTTTAAAGACCGCGCGGCCCATGAAGAGATAT 1711
QY 461 GluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGlu 480
DB 1712 GAAGAAATCAAGTATAGGTAAATGGCTTGTGTTAAGGCACAGAAAGTTCCTGAGGAA 1771
QY 481 GlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGly 500
DB 1772 GGATGGATCATGCAAGATGGCACCCATGCGCCAGGAAACAAATACCMGGGACCATCTCGGA 1831
QY 501 MetIleGlnValPheLeuGlyHisSerGlyLeuAspThrGluGlyAsnGluLeuPro 520
DB 1832 ATGATTCAGGTTTTCCTTGGTTCACAGTGTGGCTTGATCTGAGGSCCATGAGTACCC 1891
QY 521 ArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisHisLysLysAlaGly 540
DB 1892 CGTTTGTCTCTGTTCTCGTGAAGAGCGTCTGTGATTCAGCATCAAGAAAGCTGGT 1951
QY 541 AlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsn 560

1952 GCATGAATGCTCTTGTCTGTCTCAGCTGTCTTACCAATCGACAATACATGTTGAAT 2011
 561 LeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPheLeu 580
 2012 CTTGATTTGGATCACCACATACATTAACACAGTAAGGCTCTCAGGGAAGCTATGCTTCCTT 2071
 581 MetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGly 600
 2072 ATGGACCTTAACCTAGGAAGGAGTGTCTGTCTAGCTCCAGTTTCCCAGAGATTCGATGGC 2131
 601 IleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArg 620
 2132 ATTGACAGGAATGATCATGATATGCAACAGCAACACCGCTGTTTTCGATATTAACTTGACA 2191
 621 GlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArgThr 640
 2192 GGTCTTTGATGGCATCCAGGACAGTTTATGTCGGAACCTGGCTGTGTTTCAACCGAACA 2251
 641 AlaIleTyrGlyTyrGluProProIleLysAlaLysLysProGlyPheLeuAlaSerLeu 660
 2252 GCTCTATATGGTTATGAGCCCCCAATTAAAGCAGAGAAGGGTGGTTCTTGTATCATCTA 2311
 661 CysGlyGlyLysLysAlaSerLysSerLysLysArgSerSerAspLysLysSer 680
 2312 TGTGCGGTAGGAAGCAAGCAACAAATCAAGAAG---GGCTCGACAAAGAAGTCG 2368
 681 AsnLysHisValAspSerSerValProValPheAsnLeuGluAspIleGluGlyVal 700
 2369 CAGAAGCATGTGGACAGTCTGTGTCAGTATTCAACCTTGAAGATATAGAGGAGGAGTT 2428
 701 GluGlyAlaGlyPheAspAspGlyLysSerValLeuMetSerGlnMetSerLeuGluLys 720
 2429 GAAGCGCTGGATTGACGACGAGAAATCACTTCTTAATGTCTCAATGAGCTGGAGAG 2488
 721 ArgPheGlyGlnSerAlaPheValAlaSerThrLeuMetGluTyrGlyGlyValPro 740
 2489 AGATTGGCCAGTCGCGAGCGTTTGTGCTCCACTCTGATGAGTATGGTGTCTCT 2548
 741 GlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGlyTyr 760
 2549 CAGTCGCAACTCCGGAGTCTCTCTGAAGAAGACTATCCATGTTATAGCTGTGGCTAT 2608
 761 GluAspLysSerGluTyrGlyThrGluIleGlyTyrIleTyrGlySerValThrGluAsp 780
 2609 GAGGACAGACTGATGGGAACTGAGATCGGTGGATCTACGGTCTGTGACAGAAAC 2668
 781 IleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetProLys 800
 2669 ATTCTACCGGATTCAGATGACGCGAGGCTGGCGGTGCGATCTACTGCAATGCCAAG 2728
 801 ArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnValLeu 820
 2729 CGGCCAGCTTTCAAGGGGTCTGCCCCCATCAATCTTTCGACCGCTGGAACAGGTGCTC 2788
 821 ArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyrGly 840
 2789 CGGTGGCTCTTTGGTCCGTGGAGATCTCTTTCAGCCGGCACTGCCCTGTGTGTACGGC 2848
 841 TyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyrPro 860
 2849 TACGGAGGGCGGCTCAAGTCTCTGGAGAGATTCGCTATCATCAACACCATCTACCCG 2908
 861 LeuThrSerLeuProLeuValTyrCysIleLeuProAlaIleCysLeuLeuThrGly 880
 2909 CTCAGTCCATCCCGCTTCTCATCTACTGCACTCTGCCCGCATCTGTCTCACCGGA 2968
 881 LysPheIleMetProGluIleSerAsnLeuAlaSerIleTyrPheIleAlaLeuPheLeu 900
 2969 AAGTTTCATCTCCAGATCAGCAACTTCGCCAGCATCTGGTTCATCTCCCTCTTCATC 3028
 901 SerIlePheAlaThrGlyIleLeuGluMetArgTyrSerGlyValGlyIleAspGluTyr 920

Db 3029 TCGATCTTCGCCACGGGCATCTCTGGAGATGAGGTGAGCGGGGTGGGCATCGACGAGTGG 3088
 QY 921 TrpArgAsnGluGlnPheTrpValIleGlyIleSerAlaHisLeuPheAlaValPhe 940
 Db 3089 TGGAGGAACGAGCAGTTCTGGGTGATCGGGGCATCTCCGCGCACCTCTTCGCCGTTTC 3148
 QY 941 GlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLysAla 960
 Db 3149 CAGGGCCTCTCTCAAGGTGCTGGCCGCAATCGACACCACTTCACCGTCACTCCAGGCG 3208
 QY 961 AsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeuIle 980
 Db 3209 TCGGACGAGGACGGCGACTTCGCGGAGCTGTACATGTTCAAGTGGACGAGCTCTTGATC 3268
 QY 981 ProProThrThrIleLeuIleAsnMetValGlyValValAlaGlyThrSerTyrAla 1000
 Db 3269 CCGCCCAACCATCTCTGATCATCAACTGCTGGCGTCTGCGGCAAGCTCTTCTCGCTTCG 3328
 QY 1001 IleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPheTrp 1020
 Db 3329 ATCAACAGCGGATACAGTCTGGGGCCCGCTCTTCGGCAAGCTCTTCTCGCTTCG 3388
 QY 1021 ValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThrPro 1040
 Db 3389 GTCATCTCCACTGTATCCCGTCTCTCAAGGCGCTCTGCGGCGAGGAGACCGACCCCG 3448
 QY 1041 ThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpValArg 1060
 Db 3449 ACCATCGTCTGCTGCGGCATCTCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3508
 QY 1061 ValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
 Db 3509 ATCGACCCCTTCACACCCGCTCTCACTGGCCCGGATACCCAGACGTGTGGCATCACTGC 3568

RESULT 6

ADA69499

ID ADA69499 standard; DNA; 3222 BP.

XX ADA69499;

XX 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 2822.

XX Plant; bacterial infection; fungal infection; viral infection; rice;

XX gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

Identifying at least one gene involved in plant resistance or response to pathogenic infection for conferring resistance or tolerance to a plant to bacterial, fungal or viral infection by determining or detecting plant gene expression.
 Claim 6; SEQ ID NO 2822; 899pp; English.
 The present invention relates to a method (M1) for identifying genes involved in plant resistance or response to pathogenic infection. M1

CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX
 SQ Sequence 3222 BP; 805 A; 711 C; 805 G; 901 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0 Length: 3222
 Score: 5169.50 Matches: 950
 Percent Similarity: 94.52% Conservatives: 67
 Best Local Similarity: 88.29% Mismatches: 54
 Query Match: 89.47% Indels: 5
 DB: 7 Gaps: 4

US-09-900-237A-30 (1-1080) x ADA69499 (1-3222)

QY 6 AspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGlnIleCysAlaAspGly 25
 DB 4 GACGGTGCAGAAATCGGGGAAGCAA-----TGCCATGTCTGCCAGATCTCGCGCGACGGC 57
 QY 26 LeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspValCysArgPheProVal 45
 DB 58 GTGGGACGCGCGCGACGCGAGCTTTCACCGCTCGACGCTCGCGGGTCCCGGG 117
 QY 46 CysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAlaCysLeuGlnCysLys 65
 DB 118 TGCGGCGCATGCTACGAGTACGAGCGCAGGACGCGCAGCGCGTCCCGCGAGTCAAG 177
 QY 66 ThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGluGlyAspAspThr 85
 DB 178 ACCAAGTACAAGCGCACAAAGGGGAGCCCGCCGATATCTGGGGATGAAACGCGATGTT 237
 QY 86 AspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThrGluAspGlnLysGln 105
 DB 238 GATCGCGGATGATCTAGTATGATGAACTATCCAACTCCGCGCAACCGAGCCATAGCAC 297
 QY 106 LysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGlySerGlyAsnValGly 125
 DB 298 AAGATTGCCGAGAGGATGCTCACCCTGGCGCATGACTCTGGAGGAATGATATTGTC 357
 QY 126 HisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAspSerGlyGluIlePro 145
 DB 358 CATTTCTAAGTATGATAGTGTGAGATTGGTTCATCCCAAGTATGACAGTGTGAAATCCCT 417
 QY 146 ArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGluIleProGlyAlaSer 165
 DB 418 CGCATATATATCCATCCCTCCTCAGCAGCAGATCTCAGGTGAATTCCTGGAGCGTCC 477
 QY 166 ProAspHisHisMetSerProThrGlyAsnIleSerArgAla---ProPhePro 184
 DB 478 CCTGAT---CATATGATGTCCTCCGTTGGGAACATTGGCAGACGTGGGCATCCATTTCCC 534
 QY 185 TyrValAsnHisSerProAsnProSerArgGluPheSerGlySerIleGlyAsnValala 204
 DB 535 TATGTGAACCATTCACCAAAACCCATCAAGGAGTTCCTGGTAGCCCTGGCAATGTGCA 594
 QY 205 TrpLysGluArgValAspGlyTyrLysMetLysGlnAspLysGlyAlaIleProMetThr 224
 DB 595 TGGAAAGAGAGATGATGCTGGAATAATGAG---GATAAAGGTGCATTCATCCATGGCT 651
 QY 225 AsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThr 244
 DB 652 AATGCTACTAGCATTTGCTCTTCAGAAAGTCCGTGGGGTTGGTGACATTTGATGCTACT 711
 QY 245 GluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLys 264
 DB 712 GACTATAATCATGGAGCGCTTACTGAATGATGAACACGCCAACCTCTTATCGAGAAA 771

QY 265 ValProIleAlaSerSerLysIleAsnProTyrArgMetValIleValLeuArgLeuVal 284
 DB 772 GTGCCAAATTTCTGTCATCCAGATAAATCCGTACAGATGGTTATTTGTGTCGCGATGATT 831
 QY 285 ValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeu 304
 DB 832 GTCTCTATGATATTTCTGCACTACCGTATACAAATCTCTGTGGTAAATCATACCGCTA 891
 QY 305 TrpLeuLeuSerValIleCysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPhe 324
 DB 892 TGGTTGCTCTCTGTCATATGTCAGATTGGTTTGTCTATCTCTGGATCTCGGATCAGTTC 951
 QY 325 ProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAsp 344
 DB 952 CCCAAGTGGTCCCAATCAACCGTGAATCTTACCTGGATAGCTGGCTTTAAGGTATGAC 1011
 QY 345 ArgGluGlyGluProSerGlnLeuAlaAlaValAspIlePheValSerThrValAspPro 364
 DB 1012 CGGGAAGGTGAACCATCTCAATTTGGCTCTCTGTGATGATGGAGCTGCAATGTTAAATTTGAT 1071
 QY 365 LeuLysGluProProIleValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyr 384
 DB 1072 ATGAAGAACCTCTCTCTGTCATGCTCAATCTCTGCTCTCTCATCTCTGCTGGATTAC 1131
 QY 385 ProValAspLysValSerCysTyrValSerAspAspGlyAlaSerMetLeuThrPheAsp 404
 DB 1132 CCTGTTGATAAGGTATCTTGTCTATGTCATGATGGAGCTGCAATGTTAAATTTGAT 1191
 QY 405 AlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPheValLysLysTyrAsp 424
 DB 1192 GCGCTTGTGAGACTTCAGAAATTTGCTAGAAATGGTACCGTCTCTGTAAGAGTACAGC 1251
 QY 425 IleGluProArgAlaProGluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLys 444
 DB 1252 ATGAACCCAGAGCTCCAGAGTGGTACTTTGACACAGAAATTCATTACTTGAAGACAAA 1311
 QY 445 ValGlnProSerPheValLysAspArgAlaMetLysArgGlnTyrGluGluPheLys 464
 DB 1312 GTTCAGGCTCTTTTGTAAAGATCGTCGTGCTGCAAGAGGGAATATGAAGAAATTTAA 1371
 QY 465 IleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGluGluGlyTrpIleMet 484
 DB 1372 GTTCGTGTTAATGCTCTCGTCGCAAGGACACAGAAAGTCTCTGAGGAGGATGGATTATG 1431
 QY 485 GlnAspGlyThrProTrpProGlyAsnAsnThrArgAspHisProGlyMetIleGlnVal 504
 DB 1432 CAGATGGCACACTTGGCTGGTAAACAATACACAGAGACCATCTCTGGAATGATTCAGGTT 1491
 QY 505 PheLeuGlyHisSerGlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyr 524
 DB 1492 TTCCTTGGTCATAGTGGAGGCTTGTATACCAAGGCAATAGAGCTTCCTCTGTTAGTCTAT 1551
 QY 525 ValSerArgGluLysArgProGlyPheGlnHisLysLysAlaGlyAlaMetAsnAla 544
 DB 1552 GTGTCTCGTGAGAAACGCTCTGGATTCACACCATTAAGGCTGGTGTATGAATGCA 1611
 QY 545 LeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAsp 564
 DB 1612 CTGTTCTGTATCAGCTGCTCTTACTAATGGACAATCTTGTGTAATCTTGAATCTGAT 1671
 QY 565 HisTyrIleAsnAsnSerLysAlaValArgGlnAlaMetCysPheLeuMetAspProAsn 584
 DB 1672 CACTACATCAACATAGCAAGCTCTCCGAGAGGCTATGTGCTTCTCTATGATGATCAAC 1731
 QY 585 LeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAspGlyIleAspArgAsn 604
 DB 1732 CTAGGAAGGCTGTCTGTATGTCATTCATTCCTCAGAGGTTTGGCGGTATGATAGAAAT 1791
 QY 605 AspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGly 624
 DB 1792 GATCGATATGCAACACAGCAACCGTGTGTTTTCGATATTAATTTGAGAGGCTCTCGATGTT 1851
 QY 625 IleGlnGlyProValTyrValGlyThrCysValPheAsnArgThrAlaIleTyrGly 644

CC mutations in the gene, for monitoring upregulation of expression or
CC changes in enzyme activity in screening assays of compounds, for
CC detection of any number of allelic variants of the gene, or for use as
CC molecular markers in plant breeding programs. The isolated nucleic acids
CC of the present invention can also be used for recombinant expression of
CC their encoded polypeptides or for use as immunogens in the preparation
CC and/or screening of antibodies. The proteins can be employed in assays
CC for enzyme agonists or antagonists of enzyme function or for use of
CC immunogens or antigens to obtain antibodies specifically immunoreactive
CC with a protein
XX
SQ Sequence 3704 BP; 909 A; 838 C; 962 G; 994 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 0 Length: 3704
Score: 5143.00 Matches: 947
Percent Similarity: 93.99% Conservative: 69
Best Local Similarity: 87.60% Mismatches: 59
Query Match: 89.01% Indels: 6
DB: 3 Gaps: 6

US-09-900-237A-30 (1-1080) x AA299533 (1-3704)

QY	1	MetAspGlyAspAlaAspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln	20
DB	272	ATGGACGGC---GGCGACGCCACGAATTCGGGGAAGCATGTGCGCGGAGGTGTGCCAG	328
QY	21	IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal	40
DB	329	ATCTGGCGGACGGCGTGGCGACCGCGCGGAGCGGACCTCTCTCACCGCGTGGACGTC	388
QY	41	CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla	60
DB	389	TGCGGTCTCCCGTGTGGCGCCCATGCTACGAGTACGAGCGCGAAGCGCCACCGCGG	448
QY	61	CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu	80
DB	449	TGCCCGCAGTGCAGACTAAGTACAGCGCCACAAAGGAGGCCACCGATACACGGTGTAG	508
QY	81	GluGlyAspThrAspAlaAspAspGlySerAspPheAsnTyrProAlaSerGlyThr	100
DB	509	GAATATGAGGATGTGGATGTGACGATGTGAGTGTACTACAACTACCAAGCATCTGGCAAC	568
QY	101	GluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGlyGly	120
DB	569	CAGATCAGAAAGCAAAAGATTGCTGAGAGAAATGCTCACTTGGCGGCAAACTCACGTGGC	628
QY	121	SerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyrAsp	140
DB	629	AGT---GATATTGGCTGGCTAAGTATGACAGCGGTGAAATTGGGCATGGGAAGTATGAC	685
QY	141	SerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGlnMetSerGlyGlu	160
DB	686	AGTGGTGGATCCCTCGTGGATATATCCCGTCACTAACTCATACCCAGATCTCAGGAGAG	745
QY	161	IleProGlyAlaSerProAspHisHisMetSerProThrGlyAsnIleSerArgArg	180
DB	746	ATTCTCTGGAGCTCCCGCTGAT---CATATGATGTCTCTGTGGGAACATGGCAGCGGT	802
QY	181	Ala---ProPheProTyrValAsnHisSerProAsnProSerArgGluPheSerGlySer	199
DB	803	GGACATCAATTCCTTATGTAATAATCAATTCCTCAAAACCATCGAGGGAGTCTCCGGTAGC	862
QY	200	IleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLysGlnAspLysGly	219
DB	863	CTTGGCAATGTGATGGAAGAGAGGGGTGGATGGATGGAAAAATGAAG---GATAAAGGT	919
QY	220	AlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArgAlaAlaThrAsp	239
DB	920	GCAATTCCTATGACCAATGGACAGCATCTGCTCCATCAGAAAGCGGTGGAGTTGCTGAT	979
QY	240	IleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAspGluThrArgGln	259

DB	980	ATTGATGCTTCTACTGATTATAACATGGAAGTGCCTTACTCAATGATGAACCTCGCAA	1039
QY	260	ProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyrArgMetValIle	279
DB	1040	CCTCTATCTAGAAAGTCCCAATTCCTTCAATCCAGATAAATCCGTCAGAAATGGTCAAT	1099
QY	280	ValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThrAsnProValArg	299
DB	1100	GTGCTACGTTTGGCTGTCTATGTCATATTTTCCGCTACCGTATCACACATCTCTGAAAC	1159
QY	300	AsnAlaTyrProLeuTrpLeuSerValIleCysGluIleTrpPheAlaLeuSerTrp	319
DB	1160	AATGATATATCCACTGTGGCTTTTATCCGTCATATGTGAGATCTGGTTGCTTGTCTGG	1219
QY	320	IleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyrLeuAspArgLeu	339
DB	1220	ATTTTGGATCAGTTCCTCCAAAGTGGTCCCAATCAACCGTGAAACATACCTTGATAGCTG	1279
QY	340	AlaLeuArgTyrAspArgGluGlyProSerGlnLeuAlaAlaValAspIlePheVal	359
DB	1280	GCTTTAAGGTATGACCGAAGGTGAACCATCTCAATTAGCTCCTGTTGATATTTTGTCT	1339
QY	360	SerThrValAspProLeuLysGluProProIleValThrAlaAsnThrValLeuSerIle	379
DB	1340	AGTACTGTGGATCCCAATGAAAGAGCTCTCTCTGTCTACCTGCAAAATATCTGCTTCCATC	1399
QY	380	LeuAlaValAspTyrProValAspLysValSerCysTyrValSerAspAspGlyAlaSer	399
DB	1400	CTTGCTGTGATATATCCGTTTGACAAGGTATCTTGCTATGTTCGGATGATGAGGTGCT	1459
QY	400	MetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLysTrpValProPhe	419
DB	1460	ATGCTGACTTTTGTGATCTCTCTGAAACTTCAGAGTTTGTAGAAAATGGGTTCGGTTC	1519
QY	420	ValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCysGlnLysIleAsp	439
DB	1520	TGTAAGAAGTACAACTAGAGCTTANGGCCCGGATGGTACTTGTCTCAGAAAAATTGAT	1579
QY	440	TyrLeuLysAspLysValGlnProSerPheValLysAspArgAlaMetLysArgGlu	459
DB	1580	TACTTGAAGACAAAAGTTCAAACTCATTTGTGAAAGAACCGCGGCGCATGAAGAGAA	1639
QY	460	TyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeuLysValProGlu	479
DB	1640	TATGAAGAAATCAAAAGTTGATCAATGTTCTGTAGCCCAAGGCACAAAAGTTCCCGAG	1699
QY	480	GluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThrArgAspHisPro	499
DB	1700	GAGGATGGATCATGCAAGATGTTACACCTTGGCTGGGAACAATACTAGGGACCATCT	1759
QY	500	GlyMetIleGlnValPheLeuGlyHisSerGlyLeuAspThrGluGlyAsnGluLeu	519
DB	1760	GGAATGATTGAGTTTCTCGGTTCACAGTGGAGGCTTCACGTTGAAGCAATGAATCACTT	1819
QY	520	ProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHisLysLysAla	539
DB	1820	CCTCGTTGTTGTTTGTCTCGTGAACAAACGCTCTGGATTCACAAATCAACAAGAGCT	1879
QY	540	GlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGlyGlnTyrMetLeu	559
DB	1880	GGTGCCATGAATGCACTTGTTCGTGTATCAGCTGTCTTACTAATGGCAATACATGTTG	1939
QY	560	AsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGluAlaMetCysPhe	579
DB	1940	AATCTTATTTGACCACTACATCAATAATAGCAAGGCTCTTCGAGAGCTATGTGTCTTC	1999
QY	580	LeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPheProGlnArgPheAsp	599
DB	2000	CTTATGACCCCAACCTAGGAAGGAATGTCTGTTATGTTCCAAATTTCTCAGAGGTTGAT	2059
QY	600	GlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePheAspIleAsnLeu	619
DB	2060	GGTATTGATGAATGATCCGATATGCAACACAGGAACACTGTGTGTTTTCGATATTAACTTG	2119

```

QY 620 ArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCysValPheAsnArg 639
Db 2120 AGAGGCTCTGACGGCATTCAAGGCGCAGTTTATGTGGAACTGGTTGTGTGTTTAAACAGA 2179
QY 640 ThrAlaIleTyrGlyTyrGluProIleLysAlaLysLysProGlyPheLeuAlaSer 659
Db 2180 ACGGCGCTTATATGTTGTTATGAGCTCCAGTCAAGAAAAAAGCCAGGCTTCTTCTTCG 2239
QY 660 LeuCysGlyGlyLysLysLysAlaSerLysSerLysLysArgSerSerAspLysLysLys 679
Db 2240 CTTTGTGGGGAAGGAAAGAAAGCGTCAAAATCTAAGAAG--AGCTCGAAAGAAAGAGAG 2296
QY 680 SerAsnLysHisValAspSerValProValPheAsnLeuGluAspIleGlnGluGly 699
Db 2297 TCACATAGACAGCGACAGACTCTGTACAGATTATTAATCTCAAGATATAGAGAAAGG 2356
QY 700 ValGluGlyAlaGlyPheAspAspGluLysSerValLeuMetSerGlnMetSerLeuGlu 719
Db 2357 ATTGAAGGTTCTCAGTTTGTATGATGAGAAATCGCTGATTATGTCTCAATGAGCTTGGAG 2416
QY 720 LysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeuMetGluTyrGlyGlyVal 739
Db 2417 AAGAGATTGGCCGCTCAGGCTTTTGTAGGCTCTACTCTGATGGAATATGGTGGTGT 2476
QY 740 ProGlnSerSerThrProGluSerLeuLeuLysGluAlaIleHisValIleSerCysGly 759
Db 2477 CCACAATCTGCAACTCCAGAGTCTCTTCTGAAGAAGCTATTCAATGATCAGCTGTGGC 2536
QY 760 TyrGluAspLysSerGluTyrGlyThrGluLeGlyTyrIleTyrGlySerValThrGlu 779
Db 2537 TATGAGGACAAACTGACTGGGAACTGAGATTGGTGGATCTATGTTCTGTTACAGAA 2596
QY 780 AspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArgSerValTyrCysMetPro 799
Db 2597 GACATTCTCACCGGATTCAAGATGATGCTCGAGCTGGCGATCAATCTACTGCTGCT 2656
QY 800 LysArgProAlaPheLysGlySerAlaProIleAsnLeuSerAspArgLeuAsnGlnVal 819
Db 2657 AAGCGACCAAGCTTCAAGGGATCTGCTCTATCAACCTTTCGGATCGTTTGAATCAAGTG 2716
QY 820 LeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArgHisCysProLeuTyrTyr 839
Db 2717 CTTCGGTGGGCTCTTGGTTCCATTGAAATTTCTTTCAGCAGGCAATGTCCTCATATGGTAT 2776
QY 840 GlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyrIleAsnThrThrIleTyr 859
Db 2777 GGCTATGGAGCCCGCTTAAATTCCTGGAGAGATTGCTTATATCAACACACAATTTAT 2836
QY 860 ProLeuThrSerLeuProLeuLeuValTyrCysIleLeuProAlaIleCysLeuLeuThr 879
Db 2837 CCACCTCATCAATCCCGCTCTCTCTGCTACTGCATATGGCAGAGTTTGTCTTCTCACT 2896
QY 880 GlyLysPheIleMetProGluIleSerAsnLeuAlaSerIleThrPheIleAlaLeuPhe 899
Db 2897 GGGGAAGTTTCATCCCAAGAGTAGTAACCTAGAGAGTGTTCGTTTATATATGCTCTTT 2956
QY 900 LeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrPheSerGlyValGlyIleAspGlu 919
Db 2957 ATCTCAATCTTGGCATCTGGTATCCCTTGAGATGAGGTGGAGTGTGTGGCATTTGATGAA 3016
QY 920 TrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSerAlaHisLeuPheAlaVal 939
Db 3017 TGGTGGAGGAACGAGCAGTCTGGGTCAATGGTGGTATTTCTGGCAATTTATTTCGCCGTC 3076
QY 940 PheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsnPheThrValThrSerLys 959
Db 3077 TTCAGGCTCTCTGAAGGTCTTGTGTTATCGACACGAGCTTCACTGTCACCTCTAAG 3136
QY 960 AlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPheLysTrpThrThrLeuLeu 979
Db 3137 GCCACTGACGAAGAAGGTGATTTTGGCCGAGCTCTACATGTTCAAGTGGACAAAGCTTCTG 3196

```

```

QY 980 IleProProThrThrIleLeuIleLeuMetValGlyValValAlaGlyThrSerTyr 999
Db 3197 ATCCCAACCAACACATATTTTGATCATCAACCTGGTCGGCGTGGTCGTCATTTCTCTAC 3256
QY 1000 AlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGlyLysLeuPhePheAlaPhe 1019
Db 3257 GAATCAATAGCGTTACCGTCATGGGACCTCTTTTCGGGAAGCTCTCTTTTGGGTTTC 3316
QY 1020 TrpValIleValHisLeuTyrProPheLeuLysGlyLeuMetGlyArgGlnAsnArgThr 1039
Db 3317 TGGGTGATTGTCCACTGTACCCCTTCTCAAGGSCCTCATGGGAAGCAGAACCCGACG 3376
QY 1040 ProThrIleValIleValTrpAlaValLeuLeuAlaSerIlePheSerLeuLeuTrpVal 1059
Db 3377 CCGACCATGTCTGTTGTCTGGGCTATCTCTTGGTGCATCTTTTCCCTGATGTGGGT 3436
QY 1060 ArgValAspProPheThrThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsn 1079
Db 3437 CGTATCGATCCATTCCACCACCGGGTCACTGGCCCTGATATCGGAAATGTGGCATCAAC 3496
QY 1080 Cys 1080
Db 3497 TGC 3499

RESULT 8
AAC49550
ID AAC49550 standard; DNA; 3198 BP.
XX
AC AAC49550;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 61572.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-01232548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.

```

PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 13-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.

QY	296	AsnProValArgAsnAlaTyrProLeuTrpLeuSerValIleCysGluIleTrpPhe	315
Db	835	AAACCCAGTGCACAAATGCTTTGGCTCTATGGCTGGTCTCTGTGATATGTGAGATCTGGTTT	894
QY	316	AlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyr	335
Db	895	GCCTTATCCTGGATTTTGGATCAGTTTCCCAAGTGGTTTCTGTGAAACCGGTGAACACCTAC	954
QY	336	LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaAlaVal	355
Db	955	CTGCACAGGCTTGCTTTAAGATATGATCGTAGGTGAGCCATCACAGTTAGCTGCTGTT	1014
QY	356	AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr	375
Db	1015	GACATTTTCGTGAGTACTGTGTGACCCCTTGAAGGAGCCACCCCTTGTGACAGCCAAACACA	1074
QY	376	ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp	395
Db	1075	GTGCTCTCTATTCTGGCTGTGACTACCCAGTTGACAAAGGTGCTGTGTATGTTCTGTAT	1134
QY	396	AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys	415
Db	1135	GATGGTGCTGATGTTATCATTTGAAATCACTTGCAGAAACATCAGAGTTTGCCTGTA	1194
QY	416	TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435
Db	1195	TGGGTACCAATTTTGCAGAAATATAGCATAGAGCCTCGTGACCAGAAATGGTACTTTGCT	1254
QY	436	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
Db	1255	GCGAAATAGATTACTTGAAGGATAAAGTTCAGACATCAATTTGTCAAAAGATCGTAGAGCT	1314
QY	456	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
Db	1315	ATGAAGAGGGAATATGAGAAATTTAAATCCGAATCAATGCACTGTGTTCCAAAGCCCTA	1374
QY	476	LysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr	495
Db	1375	AAATGTCTGAAGAAGGGTGGTTATGCAAGATGCGACACCGTGGCTGGAAATATACA	1434
QY	496	ArgAsnHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
Db	1435	AGGACCATCCAGGAATGATCAGGTCTCTTAGGGCAAAATGGTGGACTTGTATGCAGAG	1494
QY	516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535
Db	1495	GGCAATGAGCTCCCGGTTTGGTATATGTTCTCGAGAAAGCGACCCAGGATTCACAGAC	1554
QY	536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555
Db	1555	CACAAAAAGGCTGGTGCTATGAATGCATGGTGAGAGTTTCAGCAGTTCTTACCAATGGA	1614
QY	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575
Db	1615	CCTTTCATCTTGAATCTTGGATTGTGATCATTTACATAAAATAACAGCAAAAGCCTTAAGAA	1674
QY	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595
Db	1675	GCAATGTGCTTCCTGATGGACCCAAACCTCGGAGCAAGTTTGTATGTTTCAGTTCCCA	1734
QY	596	GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe	615
Db	1735	CAAAAGATTGATGTATCGATAAAGACGATAGATATGCTAATCGTAATACCGTGTCTTT	1794
QY	616	AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys	635
Db	1795	GATATTAACTTGAGAGGTTTAGATGGGATTCAGGACCTGTATATGTCCGAACTGGATGT	1854
QY	636	ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAla-----LysLys	653
Db	1855	GTTTTCAACAGACAGCATTTATACGGTATGAACCTCCATAAAGTAAACACAGAG	1914

QY	654	ProGlyPheLeuAlaSerLeuCysGlyGlyLysLysAlaSerLysSerLysLysArg	673	Db	2995	AAAGTTGTTCTTTGGCTTCTGGGTGATTGTTCACTTGTTACCTTTCTCAAGGTTTGATG	3054
Db	1915	CAAAGTCTTTTATCTAAGCTCTGTGGTGAATCAAGAAAGAAATTCAAAGCTTAAGAAA	1974	QY	1034	GlyArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeuLeuAlaSerIle	1053
QY	674	SerSerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu	693	Db	3055	GGTCACAGAACCGACTCCACCATTTGTTGGTCTGGTCTGTTCTCTTGCTTCTATC	3114
Db	1975	GAGTCGGCAAAAAGAAATCAGCCAGGATCTACTGACTCACTGTTCTCTGTTTCAACCTC	2034	QY	1054	PheSerLeuLeuTrpValArgValAspProPheThrThrArgLeuAlaGlyProAsnIle	1073
QY	694	GluAspIleGluGluGlyValGluGlyAlaGlyPheAspAspGluLysSerValLeuMet	713	Db	3115	TTCTGTTGTTGTTGGTTAGGATTGATCCCTTCACTAGCCGAGTCACTGGCCCGACATT	3174
Db	2035	GATGACATAGAGAGAGGATTGAAGTGCTGTTTGTATGATCAAAAGCGCTCTTAATG	2094	QY	1074	GlnThrCysGlyIleAsnCys	1080
QY	714	SerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeu	733	Db	3175	CTGGAATGTGGAATCAACTGT	3195
Db	2095	TCGCAATGAGCGCTGAGAGCGATTGGACAGTCTGCTGTTTGTGCTTCTTACCTA	2154	RESULT 9			
QY	734	MetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLysGluAlaIle	753	ID	AAV06567	standard; cDNA; 3614 BP.	
Db	2155	ATGGAAATGGTGGTCTCTCTCTCAGCACTCCAGAAACCTTCTCAAGAGGCTATC	2214	XX	AAV06567;		
QY	754	HisValIleSerCysGlyTyrGluAspLysSerGluTrpGlyThrGluIleGlyTrpIle	773	XX	17-OCT-2003 (revised)		
Db	2215	CATGTCATTAGTTGTTATGAGGATAAGTCAGATTGGGGAATGGAGATTGGATGGATC	2274	DT	06-JUL-1998 (first entry)		
QY	774	TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArg	793	XX	Arabidopsis cellulose biosynthetic gene clone Ath-B.		
Db	2275	TATGGTCTGTGACAGAAATATCTGACTGGGTTTCAAAATGATGCCCTGGATGGCGA	2334	XX	Cellulose; cellulose synthase; RSW1 gene; beta-1,4-glucan;		
QY	794	SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer	813	XX	transgenic plant; ss.		
Db	2335	TCATTTTACTGCGATGCTTAGCTTCCAGCTTTCAGGGTTCGCTCTCATCACTTTTCA	2394	OS	Arabidopsis thaliana; var. Columbia.		
QY	814	AspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArg	833	XX	Key	Location/Qualifiers	
Db	2395	GATCGTCTGAACCAAGTGTGAGTGGCTTTAGGTTTCACTGAGATTCTCTTCAGTCGG	2454	FT	217.3411		
QY	834	HisCysProLeuTrpTyrGlyTyrGlyGlyArgLeuLysPheLeuGluArgPheAlaTyr	853	XX	/*tag= a		
Db	2455	CATTGCTCTATATGTTATGTTTCAATGGGAGGCTAAATTTCTTGAGAGTTTGGTAT	2514	XX	WO9800549-A1.		
QY	854	IleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuPro	873	XX	08-JAN-1998.		
Db	2515	GTGAACACCAACATCTACCTTATCCTCATCTCTCTCATGATTGTACATGGCA	2574	XX	24-JUN-1997; 97WO-AU000402.		
QY	874	AlaIleCysLeuLeuThrGlyLysPheIleMetProGluLeuSerAsnLeuAlaSerIle	893	XX	27-JUN-1996; 96AU-00000699.		
Db	2575	GCGTTTGTCTCTTCCACCACCACTTATTTATCTCAGATTAGTAACTATGCAAGTATA	2634	XX	(AUS) UNIV AUSTRALIAN NAT.		
QY	894	TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSer	913	XX	(CSIR) COMMONWEALTH SCI & IND RES ORG.		
Db	2635	TGGTTTCTGCTCTCTTCTCTCTTTCATTTTCCACCGGATATCTAGAAATGAGTGGAGT	2694	XX	Arioli A, Williamson RE, Betzner AS, Peng L;		
QY	914	GlyValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSer	933	XX	WPI; 1998-086974/08.		
Db	2695	GCGGTAGCATAGACGATGTGGAGAAACGAGCAGTTTGGGTCAATGGTGGAGTATCC	2754	XX	P-PSDB; AAW33819.		
QY	934	AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn	953	XX	DNA encoding cellulose biosynthetic enzyme - useful for manipulation of		
Db	2755	GCTCATTTATTCGCTGTCTTCAAGGTATCTCCTCAAGTCCCTTGGCGGTATGGACACAAC	2814	XX	cellulose and beta-1,4-glucan.		
QY	954	PheThrValThrSerLysAlaAsnAspGluGlyAspPheAlaGluLeuTyrMetPhe	973	XX	Claim 12; Page 136-143; 207pp; English.		
Db	2815	TTACAGATTACTCAAAAGCTTCAGATGACGAGGACTTTGCTGAGCTTCTACTTGTTC	2874	XX	cDNA clone Ath-B was isolated from an Arabidopsis thaliana cDNA library		
QY	974	LysTrpThrThrLeuIleProProThrThrIleLeuIleIleAsnMetValGlyVal	993	XX	using probes obtained by PCR with primers (see AAT99632-34) based on		
Db	2875	AAATGGACAACACTTCTGATTCGCGCAACGAGCTGCTCATTTGTAATCTAGTGGAGT	2934	XX	cellulose synthase RSW1 genomic clone 23H12 (see AAV06563) and EST clone		
QY	994	VallAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGly	1013	XX	AA120782 (see AAV06562). It is closely related to RSW1 cDNA (see		
Db	2935	GTTGAGGAGTCTCTTATGCTATCAACAGTGGATACCAATCATGCGGACCACTTTTGT	2994	XX	AAV06565) and to partial genomic clone 12C4 (see AAV06564), and encodes a		
QY	1014	LysLeuPheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMet	1033	XX	1065-polypeptide (see AAW33819). Claimed nucleic acid molecules (see		
				XX	AAV06562-69) coding for claimed polypeptides (see AAW33816-20 and		
				XX	AAW46202) involved in cellulose biosynthesis can be used to manipulate		
				XX	the cellulose and/or beta-glucan content of transgenic plants. Expression		
				XX	of nucleic acids in the sense orientation increases the level of		
				XX	cellulose and reduces the level of non-crystalline beta-1,4-glucan and		
				XX	starch, providing plants with modified strength and/or shape and/or fibre		
				XX	properties, or having increased resistance to stresses or pests.		
				XX	Antisense, ribozyme or co-suppression molecules can be used to reduce the		
				XX	cellulose content of a transgenic plant, e.g. to improve digestibility or		
				XX	to alter carbon partitioning such that increased carbon is available for		
				XX	growth, rather than deposited as cellulose. (Updated on 17-OCT-2003 to		

```
CC standardise OS field)
XX
SQ Sequence 3614 BP; 963 A; 717 C; 863 G; 1071 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3614
Score: 4499.00 Matches: 835
Percent Similarity: 87.44% Conservative: 98
Best Local Similarity: 78.26% Mismatches: 108
Query Match: 77.86% Indels: 26
DB: 2 Gaps: 8

US-09-900-237a-30 (1-1080) x AAV06567 (1-3614)

QY 19 CysGlnIleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCys 38
Db 274 TGCAGATCTGTAGTACAAATGTTGCGACAGCTGTGTAGTGGAGATCGTTTGTGGCTTGT 333
QY 39 AspValCysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThr 58
Db 334 GATATTGTGTATCTCCAGTTTGTGCGCTTGTCTACGAGTATGAGAGGAAGATGGGAAT 393
QY 59 GlnAlaCysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArg 78
Db 394 CAATCTGTCTCAGTGCAGAAACCAAGATCAAGAGGCTCAAGAGGTAGTCTGTATTCCT 453
QY 79 GlyGluGlyAspAspThrAspAlaAspAspGlySer--AspPheAsnTyrProAla 97
Db 454 GGTGATGAAGACGAGGATGGCTTAGCTGATGAGGTACTGTGAGTTCAACTACCCCT--- 510
QY 98 SerGlyThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsn 117
Db 511 -----CAGAAGGAGAAATTTTCAGAGCGGATCGTTGGTTGGCATCTTACT 555
QY 118 ThrGlyGlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSer 137
Db 556 CGTGGGAAGGAGAGAAATGGGGAAACCCAGTATGATAA----- 597
QY 138 LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsnSerGln--- 156
Db 598 -----GAGGTCTCTCACAATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 642
QY 157 MetSerGlyGluIleProGlyAlaSerProAspHisHisMetSerProThrGlyAsn 176
Db 643 ACTTCAGAGAGATTTCTGCTGCTCCTCACCTGAAGCGCTCTCTGTATCTTCTACTATCGT 702
QY 177 IleSerArgAlaProPheProTyr--ValAsnHisSerProAsnProSerArgGlu 195
Db 703 GGGGGAAGCGCTTCTCTATCTATCATAGATGCAATCAATCACCAMAATAGAAGGATTGTG 762
QY 196 PheSerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLys 215
Db 763 GATCCTGTTGCGACTCGGAATGTAGCTTGAAGAGAGAGTGTATGCTGGAAATGAAG 822
QY 216 GlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg 235
Db 823 CAAGAGAAGAATACTGCTCTCTGTC-----AGCAGCGAGGCTGCTTCTGAA--AGA 870
QY 236 AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp 255
Db 871 GGTGGAGTAGATATTGATGCCAGCACAGATATCTAGCAGATAGGCTCTGCTGAATGAC 930
QY 256 GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr 275
Db 931 GAAGCGAGGCGAGCTTCTGTCAAGGAAGTTTCAATCTCTTCATCACGGATCAATCCCTTAC 990
QY 276 ArgMetValIleValLeuArgLeuValLeuValSerIlePheLeuHisTyrArgLeuThr 295
Db 991 AGAATGGTATTATGTCGCGCTTGTATCTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1050
QY 296 AsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPhe 315
Db 1051 AACCCAGTGCCAAATGCCCTTGTCTATGCTGGTCTGTGTATATGTGAGATCTGGTTT 1110
```

```
QY 316 AlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyr 335
Db 1111 GCCTTATCTCTGGATTTTGGATCAGTTTCCCAAGTGGTTTCTGTGAACCGGTGAACCTAC 1170
QY 336 LeuAspArgGluAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaVal 355
Db 1171 CTCGACAGGCTTGTCTTAAAGATATGATCGTGAAGGTGAGCCATCACAGTTAGTCTGTT 1230
QY 356 AspIlePheValSerThrValAspProLeuLysGluProIleValThrAlaAsnThr 375
Db 1231 GACATTTTCTGTGAGTACTGTGACCTTGAAGAGGCCCTTGTGACAGCCACACACA 1290
QY 376 ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp 395
Db 1291 GTGCTCTCTATCTGTGCTGTGTGACTACCCAGTTGACAGGTGCTCTGTATGTTCGTAT 1350
QY 396 AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys 415
Db 1351 GATGGTGTCTGTATGTATCATTTGAATCACTTGCAGAAACATCAGAGTTTGTCTCTGATA 1410
QY 416 TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys 435
Db 1411 TGGGTACCATTTTTCAGAAATATATGATAGAGCCCTCGTGCACCAAGATGTACTTTGCT 1470
QY 436 GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla 455
Db 1471 GCGAAATATAGATTACTTGAAGGATTAAGTTCAGACATCATTTGTCAAGATCGTAGAGCT 1530
QY 456 MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu 475
Db 1531 ATGAAGAGGGAATATGAGGAATTTAAATCCGAATCAATGCACTTGTTCCTCAAGCCCTA 1590
QY 476 LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr 495
Db 1591 AAATGCTCTGAGAAGGGTGGTTAAGCAAGATGCAACCGTGGCTGGAAATAATACA 1650
QY 496 ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyLeuAspThrGlu 515
Db 1651 GGGGACCATTCCAGGAATGATCCAGGTCTTCTTAGGGCAAAATGGTGGCTGTATCCAGAG 1710
QY 516 GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis 535
Db 1711 GGCAATGAGCTCCCGCTTGTGTATATGTTTCTCGAGAAAAGCAGCAGGATTCAGCAC 1770
QY 536 HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly 555
Db 1771 CACAAAAGGCTGGTGTATGAATGCACTGGTGAGAGTTTCAGCAGTTCTTTACCAATGGA 1830
QY 556 GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu 575
Db 1831 CCTTTCATCTTGAATCTTGTATGTGATCAATACATAATAACCAAGCCCTTAAGAGAA 1890
QY 576 AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro 595
Db 1891 GCAATGGCTTCTGTATGAGCCCAACCTCGGAAAGCAAGTTTGTATGTTCAGTTCCCA 1950
QY 596 GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe 615
Db 1951 CAAAGATTGATGGTATCGATAGAACGATAGATATGCTAATCGTAAATCCGTTGTTCTT 2010
QY 616 AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys 635
Db 2011 GATATTAACTGAGAGGTTTAGATGGATTCAGGACCTGTATATGTTCGGAACCTGATGT 2070
QY 636 ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAla-----LysLys 653
Db 2071 GTTTTCAACAGAACACCATATACGGTTATGAACCTTCAATTAAGTAAACACACAGAG 2130
QY 654 ProGlyPheLeuAsnLeuSerCysGlyGlyLysLysLysAlaSerLysSerLysLysArg 673
Db 2131 CCAAGTCTTTATCTAAGCTCTGTGTGTGATCAAGAAAGAAATTCACCAAGCTAAGAAA 2190
```

QY	674	SerSerAspLysLysLysSerAsnLysHisValAspSerValProValPheAsnLeu	693
Db	2191	GAGTCGGACAAAGAATAATCAGGACGCATCTACTCACTGTTCTGTATTCAACCTC	2250
QY	694	GluaSpilleGluGluGlyValGluGlyAlaGlyPheAspAspGluLysSerValLeuMet	713
Db	2251	GATCATAGGAAGAGGAGTGAAGTGCTGGTTTGTATGATGAAAAAGCGCTTAATG	2310
QY	714	SerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeu	733
Db	2311	TGCAAAATGAGCTGGAGAACGATTTGGACAGTCTGCTGTTTTGTGTTCTTACCCCTA	2370
QY	734	MetGluTyrglyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIle	753
Db	2371	ATGGAAAATGGTGGTGTCTCTCTTCAGCACTCCAGAAAACCTTCACAAAGAGGCTATC	2430
QY	754	HisValIleSerCysGlyTyrgluAspLysSerGluTrpGlyThrGluIleGlyTrpIle	773
Db	2431	CATGTCATAGTTGTTGTTATGAGGATAAGTCAGATTGGGAATGGAGATTGGATGGATC	2490
QY	774	TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTrpArg	793
Db	2491	TATGCTTCTGTACAGAGATATTCTGCTGGTTCAAAATGCATGCCCGTGATGGCGA	2550
QY	794	SerValTyrcysMetProlysArgProAlaPheLysGlySerAlaProIleAsnLeuSer	813
Db	2551	TCCATTATCTGATGCTTAAGCTTCAGGTTTCAAGGTTCTGCTCCTCAATCAATCTTCA	2610
QY	814	AspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArg	833
Db	2611	GATGCTGACCAAGTCTGAGTGGCTTTAGGTTTCAGTTTCAGATTCTCTTCAGTCGG	2670
QY	834	HisCysProLeuTrpTyrGlyTyrglyArgLeuLysPheLeuGluArgPheAlaTyrr	853
Db	2671	CATTGCTCATATGTTATGTTTACAATGGAGGCTAAAATTTCTTGAGAGGTTTGGCTAT	2730
QY	854	IleAsnThrThrlleTyrrProLeuThrSerLeuProLeuValTyrcysIleLeuPro	873
Db	2731	GTGAACACCACCATCTACCTATACCTATCCATCTCTCTCATGTATTGTACATTGCTA	2790
QY	874	AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle	893
Db	2791	GCGTTTGTCTCTTCCACCAACAGTTTATTATTCTCAGATTAGTAACATTGCAAGTATA	2850
QY	894	TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSer	913
Db	2851	TGGTTTCTGCTCTCTCTCTCCATTTTCGCCACGGGTATCTAGAAATGAGGTGAGT	2910
QY	914	GlyValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyIleSer	933
Db	2911	GGCGTAGGCTATAGCAAGTGGTGAGAAACAGACAGTTTGGGTTCATTGGTGAGTATCC	2970
QY	934	AlaHisLeuPheAlaValPheGlnGlyLeuLysValLeuAlaGlyIleAspThrAsn	953
Db	2971	GCTCATTTATCGCTGTGTTTCAAGGTATCTCAAAGTCTCTGCGGTATTGACACAAAC	3030
QY	954	PheThrValThrSerLysAlaAsnAspGluGlyAspPheAlaGluLeuTyrrMetPhe	973
Db	3031	TTCAAGTTACCTCAAAGCTTCAGATGAAGACGAGACTTTGCTGAGCTCTACTTGTTC	3090
QY	974	LysTrpThrThrLeuLeuIleProProThrThrIleLeuIleIleAsnMetValGlyVal	993
Db	3091	AAATGGACACACTCTGATTCGCCCAACGACGCTGCTCATGTAAACTTAGTGGAGTT	3150
QY	994	ValAlaGlyThrSerTyrrAlaIleAsnSerGlyTyrrGlnSerTrpGlyProLeuPheGly	1013
Db	3151	GTTGAGAGGCTCTTATGCTATCAACAGTGTGATACCAATCATGGGACCACTCTTGGT	3210
QY	1014	LysLeuPhePheAlaPheTrpValIleValHisLeuTyrrProPheLeuLysGlyLeuMet	1033
pB	3211	AAGTTGTTCTTGGCTTCTGGGTGATTGTTCTACTTGTACCTTCTCCTCAAGGTTGATG	3270
QY	1034	GlyArgGlnAsnArgThrProThrIleValIleValTrpAlaValLeuAlaSerIle	1053

CC with a protein

XX Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 U; 0 Other;
 Alignment Scores:

Pred. No.: 0 Length: 3725
 Score: 3993.50 Matches: 749
 Percent Similarity: 80.64% Conservative: 134
 Best Local Similarity: 68.40% Mismatches: 153
 Query Match: 69.12% Indels: 59
 DB: 3 Gaps: 20

US-09-900-237A-30 (1-1080) x AAZ99500 (1-3725)

QY 2 AspGlyAspAlaAspAla---LeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
 DB 245 GACGGCGAGCCCTGTCCGGCTAAGCCACGAGAGTGCGAATGGGCGAGGTCTGCCAG 304
 QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
 DB 305 ATTTGTGGCGACACTGTTGGCGTTTCAGCCACTGGTGAATGCTTGTGTCCTGCAATGAG 364
 QY 41 CysArgPheProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
 DB 365 TGTGCTTCCCTGCTGCGCGCCCTTGTCTATGATGATGATGATGATGATGATGATGATGAT 364
 QY 61 CysLeuGlnCysLysThrLysTyrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
 DB 425 TGCCCTCAGTCAAGACTAGATACAGAGACAGAAAGGTAGCCCTCGAGTTTCATGGTGTAT 484
 QY 81 GluGlyAspAspThrAspAlaAspAsp---GlySerAspPheAsnTyrProAlaSerGly 99
 DB 485 GAT---GAGGAGGAAGATGTTGATGACCTGGGACAATGAATCAACTAT-----529
 QY 100 ThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTyrArgMetAsnThrGly 119
 DB 530 -----AAGCAA-----535
 QY 120 GlySerGlyAsnValGlyHisProLysTyr-----AspSerGlyGluIleGly 135
 DB 536 -----GSCATGGGAAGGGCCAGAGTGGCAGCTTCAAGGAGATGACGTGATCTGTCT 589
 QY 136 LeuSer---LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsn 154
 DB 590 TCATCTGTCTGCCATGAC-----CCACACCATCGAATTCACGCGCTTACAACT 637
 QY 155 --SerGlnMetSerGlyGluIleProGlyAlaSerProAspHisHis---MetMetSer 172
 DB 638 GGACAAACAGATATCTGGAGAGATCCCTGATCCCTGACCTCCCGTCACTTCTATCCGCACT 697
 QY 173 ProThrGlyAsn---IleSerArgArgAlaProPheProTyrValAsnHisSerProAsn 191
 DB 698 CCAACATCGAGCTATGTTGATCCAAAGCGTTCAGTTCCTGTG-----AGGATTTGGGAC 751
 QY 192 ProSerArgGluPhe---SerGlySerIleGlyAsnValAlaIleTyrLysGluArgValAsp 210
 DB 752 CCTCTGAAGACTTGAATTCCTATGGGCTTAATAGTGTGATCGAAGGAAAGATTGAG 811
 QY 211 GlyTyrLysMetLysGlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAla 230
 DB 812 AGCTGGAGGGTTAAACAGGACAAATAATGTTGCAAGTGACTAAT-----856
 QY 231 ProSerGluGlyArgAlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAsp 250
 DB 857 ---AAATATCCAGGGCTAGAGGAGACATGGAG---GGGACTGGCTCAAAATGGAGAAGAT 910
 QY 251 AlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLysValProIleAlaSerSer 270
 DB 911 ATGCAAAATGGTTGATGATGACAGCGCTACCTTGGCCGCGATTTGGCCAAATTTCTCAAAC 970
 QY 271 LysIleAsnProTyrArgMetValIleValLeuArgLeuValValLeuSerIlePheLeu 290
 DB 971 CAGCTCAACCTTTTACCGGATAGTAATCATCTCCGTTTATCATCTCTGCTGCTCTCTCTTC 1030

QY 291 HisTyrArgLeuThrAsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIle 310
 DB 1031 CAATATCGTATCAGTCACTCAGTCGCTAATGCTTATGGATTGGCTAGTATCTGTATC 1090
 QY 311 CysGluIleTyrPheAlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIle 330
 DB 1091 TGTGAGTCTGGTTTGGCTTGTCTGGCTTCTAGATCAGTTCCTCAAAATGGTATCCAATC 1150
 QY 331 AsnArgGluThrTyrTyrLeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSer 350
 DB 1151 AACCGTGAGACATACTCGACAGGCTTGCATGAGGTATGATAGAGGAGGAGCCATCA 1210
 QY 351 GlnLeuAlaAlaValAspIlePheValSerThrValAspProLeuLysGluProIle 370
 DB 1211 CAGCTGGCTCCCATTTGATGATGCTTTGTCCATTCTGTGTGGATTACCCCTGTTGACAAAGTGCA 1270
 QY 371 ValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSer 390
 DB 1271 ATCAGACGCCACACACTGTTTGTCCATTCTGTGTGGATTACCCCTGTTGACAAAGTGCA 1330
 QY 391 CysTyrValSerAspAspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSer 410
 DB 1331 TGCTATGTTTCTGATGATGGCTCAGCTATGCTGACTTTTGAGTCTCTCTGAAACTGCC 1390
 QY 411 GluPheAlaArgLysTyrValProPheValLysLysTyrAspIleGluProArgAlaPro 430
 DB 1391 GAATTTGCTAGAAAGTGGGTTCCTTTTGAAGAAGCACAAATATTGAACCAAGAGCTCCA 1450
 QY 431 GluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGlnProSerPheVal 450
 DB 1451 GAATTTTACTTTGCTCAAAAATAGATTACCTCGAAGGACAAATATTCAATCCCTTCAATTTGTT 1510
 QY 451 LysAspArgArgAlaMetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeu 470
 DB 1511 AAGGMAAGACCGAGCNAATGAAGAGAGAGTATGAGAAATTCAAAATAAGAAATCAATGCCCTT 1570
 QY 471 ValSerLysAlaLeuLysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrp 490
 DB 1571 GTTGCCAAAGCACAGAAAGTGCCTGAAGAGGGTGGACCATGGCTGATGGAACCTGCTGG 1630
 QY 491 ProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGly 510
 DB 1631 CCTGGGAATAACCTTAGGACCATCTCGGCATGATTCAGGTGTCTTCTGGGCGACAGTGGT 1690
 QY 511 GlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArg 530
 DB 1691 GGGCTTGACACTGATGGAATGAATTAACACAGTCTTGTCTATGCTCTCTGCGGAAAGAGA 1750
 QY 531 ProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAla 550
 DB 1751 CCAGGCTTTGAGCATCAAGAGGCTGGTGCATGAATGAATGACCTGATTCGTGTATCTGCT 1810
 QY 551 ValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSer 570
 DB 1811 GTGCTGCAAAATGCTGCTATCTTCTCAATGTGGATTGTGACCATTAATCAATAGCAGC 1870
 QY 571 LysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCys 590
 DB 1871 AAAGCTCTTAGAAGAGCAATGTGCTTCATGATGGATCCAGCTCTAGGAAGGAAACTTGT 1930
 QY 591 TyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArg 610
 DB 1931 TATGTCAAAATTTCCACAAAGATTGTGATGGCATTTGACCTGACGATCGATGCTAATAGG 1990
 QY 611 AsnThrValPhePheAspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyr 630
 DB 1991 AACATAGTCTTCTTGTATATCAACATGAAGGCTGTAGATGGCATTCAGGCTCCAGTCTAT 2050
 QY 631 ValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLys 650
 DB 2051 GTGGGAACAGGATGCTGTTCAATAGCAGGCTTTGTATGGATGATGATGCTCTTTTGA 2110

QY	651	AlaLys-----LysProGlyPheLeuAlaSerLeuCysGlyGlyGlyLeuLysAlaSer	668
DB	2111	GAAGCTGATCTGGAACCTAACATTGTTTAAGAGCTGCTGTGTAGAGGAAG---AGA	2167
QY	669	LysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisValasSerSerVal	688
DB	2168	AAGAACAAGAGTTATATGATGATGCTCAAGCCGCTATTATCAAGAGAACAGAAATCTTCAGCT	2227
QY	689	ProValPheAsnLeuGluAspIleGluGluGlyValGluGlyAlaGlyPheAspAspGlu	708
DB	2228	CCCATCTTTTAACATGGAAGACATCGAGGAGGGTATTGAA-----GGTTATGAGAGTAA	2281
QY	709	LysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPhe	728
DB	2282	AGGTCAGTGTATGTCCTCCAGAGAAATTTGGAGAAACGCTTTGGTCAGCTCCCAATCTTC	2341
QY	729	ValAlaSerThrLeuMetGluTyrGlyGlyValProGlnSerSerSerThrProGluSerLeu	748
DB	2342	ATTGCATCCACCTTTATGACTCAAGGTGGGCATACCACTTCAACAAACCCAGCTTCTCTA	2401
QY	749	LeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSerSerGluTyrGlyThr	768
DB	2402	CTGAAGGAAGCATATCCATGTTATCAGCTGTGGGTACAGGACAAACTGAATGGGAAAA	2461
QY	769	GluIleGlyTyrIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHis	788
DB	2462	GAGATTGGCTGGATCTATGTTTCAGTTTACAGAGGATATTCTGACTGGGTTTAAAAATGCAT	2521
QY	789	AlaArgGlyTyrArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAla	808
DB	2522	GCAAGAGGCTGGCAATCAATCTACTGCATGCCACCAAGACCTTGTTCGAAGGGTCTGCA	2581
QY	809	ProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGlu	828
DB	2582	CCAACTCAATCTTCTGATGCTCTTAATCAGTGCTCCGTGGGCTCTTGGTCAGTGGAA	2641
QY	829	IleLeuPheSerArgHisCysProLeuTyrTyrGlyTyrGlyArgLeuLysPheLeu	848
DB	2642	ATTCTGCTTAGCAGACATTGCTCTATATGTTATGGCTTACAATGGCGAATGAAGCTTTTG	2701
QY	849	GluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuVal	868
DB	2702	GAGAGGCTGGGCTTACATTAACACCATTTGTTATCCAATCACTCTGTTCGGCTTATCGGC	2761
QY	869	TyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSer	888
DB	2762	TATTTGTGCTTCCGCTATCTGCTCTTACCAATAAATTAATCATCTCCTGAGATTAGT	2821
QY	889	AsnLeuAlaSerIleTyrPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeu	908
DB	2822	AAATTATGCTGGAAATGTTCTTCATTCTTTTGGCTCCCATTTTCGCACTGATATTATG	2881
QY	909	GluMetArgTyrSerGlyValGlyIleAspGluTyrTyrArgAsnGluGlnPheTyrVal	928
DB	2882	GAGCTCAGATGAGTGGTGTGGCAATGGTGGAATGGTGGAATAAGACAGTTCCTGGCTT	2941
QY	929	IleGlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAla	948
DB	2942	ATTGTGGGCACCTTCGCCCATCTCTTCGGGGTGTCCAGGGCTCTGCTGAAAGTGTGGCT	3001
QY	949	GlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAla	968
DB	3002	GGGATTGATACCAACTTCAGCTTACCTCAAAGGCATCTGATGAGGATGGCAGCTTGTCT	3061
QY	969	GluLeuTyrMetPheLysThrThrLeuLeuIleProProThrThrIleLeuIleIle	988
DB	3062	GAGCTATATGTTTCAAGTGGACCAAGTTTGCTTCATCCCTCCGACCACTGTTCTGTGCTAT	3121
QY	989	AsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyr	1008
DB	3122	AACCTGGTCGGAATGTTGGTCAGGAATTCGTATGCCATTAAACAGCGCTACCAATCTCTGG	3181
QY	1009	GlyProLeuPheGlyLysLeuPhePheAlaPheTyrValIleValHisLeuTyrProPhe	1028

[illegible]

Qy	631	ValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTrpGluProPheLeuLys	650
Db	2051	GTGGAAACAGAGATGCTGTTTCAATAGGAGGCTTTGTATGGATATGATCCTGTTTGACT	2110
Qy	651	AlaLys-----LysProGlyPheLeuAlaSerLeuCysGlyGlyLysLysAlaSer	668
Db	2111	GAAGCTGATCTGGAACCTTAACATTGTTTAAAGAGCTGCTGGTAGAAGGAAG--AGA	2167
Qy	669	LysSerLysLysArgSerSerAspLysLysSerAsnLysHisValAspSerSerVal	688
Db	2168	AAGAACAAAGAGTTATATGATGATAGTCAAAGCCGTAATTATGAGAGAACAGAACTTC	2227
Qy	689	ProValPheAsnLeuGluAspGluGlyValGluGlyValGluGlyPheAspGlu	708
Db	2228	CCCATCTTTAAATGAAGACATCGAGAGGGGTAATTGAA-----GGTTATGAGATGAA	2281
Qy	709	LysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPhe	728
Db	2282	AGGTGAGTGTATGTCACAGAGAAATTTGGAGAAACGCTTTGTCAGTCTCAATCTTC	2341
Qy	729	ValAlaSerThrLeuMetGluTyrGlyValProGlnSerSerThrProGluSerLeu	748
Db	2342	ATTGATCCACCTTTATGACTCAAGGTGGCATAACCACTTCAACAAACCCAGCTTCTCTA	2401
Qy	749	LeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSerGluTrpGlyThr	768
Db	2402	CTGAAGGAGAGCTTATCATGTTATCAGCTGTGGGTACGAGACAAACTGATGGGGAAA	2461
Qy	769	GluIleGlyTrpIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHis	788
Db	2462	GAGATTGGCTGATCATGTTTACAGTACAGAGATATTCTGACTGGGTTTAAATGTCAT	2521
Qy	789	AlaArgGlyTrpArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAla	808
Db	2522	GCAAGAGGCTGGCAATCAATCTACTGCATGCCACCAAGCTTTGTTCAAGGGTCTTGCA	2581
Qy	809	ProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGlu	828
Db	2582	CCAATCAATCTTCTGATGCTTAATCAGTGCTCGTTGGGCTCTTGGGTCTAGTGGAA	2641
Qy	829	IleLeuPheSerArgHisCysProLeuTrpTyrGlyTyrGlyArgLeuLysPheLeu	848
Db	2642	ATTCTGCTAGCAGACATTCCTATATGATGCTGCTACAAATGGCGGATTTGAAGCTTTTG	2701
Qy	849	GluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuVal	868
Db	2702	GAGAGGCTGGCTTACATTAACACCATTTGTTTATCCATCATCATCTGTTCCGCTTATGCC	2761
Qy	869	TyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSer	888
Db	2762	TATTGTGTGCTTCTGCTATCTGTTCTTCTTACCAATTAATTTATCTCTGAGATTAGT	2821
Qy	889	AsnLeuAlaSerIleTrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeu	908
Db	2822	AATTATGCTGGAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2881
Qy	909	GluMetArgTrpSerGlyValGlyIleAspGluTrpArgAsnGluGlnPheTrpVal	928
Db	2882	GAGCTCAGATGGAGTGTGTTGGCATTTGAAGATTGGTGGAGAAATGAGCAGTTTGGGT	2941
Qy	929	IleGlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAla	948
Db	2942	ATTGTTGGACCTTGTGCCATCTCTTGTGGGTGTTCCAGGGTCTGCTGAAAGTGTGGCT	3001
Qy	949	GlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAla	968
Db	3002	GGGATTGATACCACTTACAGTTTACCTCAAAGGCATCTGATGAGGATGGCGACTTTGCT	3061
Qy	969	GluLeuTyrMetPheLysTrpThrLeuLeuIleProThrThrIleLeuIleIle	988
Db	3062	GAGCTATATGTTTCAAGTGACCATGTTGCTTCATCTCCCTCCGACCACTGTTCTTGTCAT	3121
Qy	989	AsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrp	1008

Db 3122 AACCTGGTCGGAATGTCGCGAATTTTCGATGTCATTAACAGCGGCTACCAATCTCTGG 3181

Qy 1009 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe 1028

Db 3182 GGTCCGCTCTTTGGAAAGCTGTTCTTCGATCTGGGTGATCTCTCACTCTTACCCCTTC 3241

Qy 1029 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleValTrpAlaVal 1048

Db 3242 CTCAGGGTCTCATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3301

Qy 1049 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgValAspProphe-----Thr 1065

Db 3302 CTCCTTGGCTCTATCTTCTCTTGTGGGTGAGATCGATCTTTCATCTCCCGACA 3361

Qy 1066 ThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080

Db 3362 CAGAAAGCTGCCG-----TTGGGGCAATGTTGGTGAAGTGC 3400

RESULT 12

AAZ99506

ID AAZ99506 standard; DNA; 3725 BP.

XX AAZ99506;

AC AAZ99506;

XX 03-JUL-2000 (first entry)

XX DNA encoding a maize cellulose synthase.

XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;

KW transgenic plant; plant breeding marker; ss.

XX Zea mays.

XX Key Location/Qualifiers

PH CDS 179..3397

FT /*tag= a

FT /product= "cellulose synthase"

XX /note= "no termination codon given"

PN WO200009706-A2.

XX

XX 24-FEB-2000.

XX

XX 16-AUG-1999; 99WO-US018760.

XX

XX 17-AUG-1998; 98US-0096822P.

XX

XX (PION-) PIONEER HI-BRED INT INC.

XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX WPI; 2000-224343/19.

XX P-PSDB; AAY84112.

XX

XX New genes which encode maize cellulose synthase polypeptides in plants

XX useful for modulating the expression of cellulose synthase in plants and

XX to produce transgenic plants expressing the novel protein.

XX Claim 1; Page 121-126; 119pp; English.

XX

XX The present sequence encodes a maize cellulose synthase polypeptide. The

XX cellulose synthase can be used for the improvement of stalk quality for

XX improved stand or silage. It also provides an increased concentration of

XX cellulose in the pericarp, hardening the kernel and improving its

XX handling ability. The sequences are used to produce transgenic plants and

XX seeds expressing the cellulose synthase. The polynucleotide is used for

XX modulating, preferably increasing, the level of the synthase in a plant

XX cell. The plants are preferably monocots. The polynucleotide is also used

XX as a probe or primer in the detection quantitation or isolation of gene

XX transcripts. The probes are useful in detecting deficiencies in the level

XX of mRNA in screenings for desired transgenic plant, for detecting or

XX mutations in the gene, for monitoring upregulation of expression or

CC changes in enzyme activity in screening assays of compounds, for
 CC detection of any number of allelic variants of the gene, or for use as
 CC molecular markers in plant breeding programs. The isolated nucleic acids
 CC of the present invention can also be used for recombinant expression of
 CC their encoded polypeptides or for use as immunogens in the preparation
 CC and/or screening of antibodies. The proteins can be employed in assays
 CC for enzyme agonists or antagonists of enzyme function or for use of
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with a protein
 XX

SQ Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3725
 Score: 3993.50 Matches: 749
 Percent Similarity: 80.64% Conservative: 134
 Best Local Similarity: 68.40% Mismatches: 153
 Query Match: 69.12% Indels: 59
 DB: 3 Gaps: 20

US-09-900-237A-30 (1-1080) x AAZ99506 (1-3725)

QY 2 AspGlyAspAlaaspala---LeuLysSerGlyArgHisGlyAlaGlyAspValCysGln 20
 DB 245 GACGGCGAGCGCTGTCCTCCGGCTAAGCCCAAGAGAGTGCAGATGGCAGGTCTGCCAG 304
 QY 21 IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal 40
 DB 305 ATTTGTGGCGACACTGTTGGCGTTTCAGCCACTGGTGATGCTTTGTGCTCCATGAG 364
 QY 41 CysArgPheProValCysArgProCystyrGluHisGluArgLysGluGlyThrGlnAla 60
 DB 365 TGTGCTCTCCCTGCTCTGGCCCTTGTCTATGATGACGAGCGCAAGAGGGCAACATGTC 424
 QY 61 CysLeuGlnCysLysThrLysThrLysArgHisArgGlySerProAlaIleArgGlyGlu 80
 DB 425 TGCCCTCAGTGCAGACTAGATACAGAGACAGAAAGGTAGCCCTCGAGTTTCATGCTGAT 484
 QY 81 GluGlyAspAspThrAspAlaAspAsp---GlySerAspPheAsnTyrProAlaSerGly 99
 DB 485 GAT---GAGGAGGAGATGTTGATGACTGCAATGAATTCACATAT-----529
 QY 100 ThrGluAspGlnLysLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGly 119
 DB 530 -----AAGCAA-----535
 QY 120 GlySerGlyAsnValGlyHisProLysTyr-----AspSerGlyGluIleGly 135
 DB 536 -----GCCAATGGGAAGGGCCAGAGTGGCAGCTTCAAGGAGATGACGCTCATCTGTCT 589
 QY 136 LeuSer---LysTyrAspSerGlyGluIleProArgGlyTyrValProSerValThrAsn 154
 DB 590 TCATCTGCTGCCATGAC-----CCACACCATCGGATTCACGCTTCATCAAGT 637
 QY 155 ---SerGlnMetSerGlyGluIleProGlyAlaSerProAspHisHis---MetMetSer 172
 DB 638 GGCAACAGAGATATCTGGAGAGATCCCTGATCATCCCTGACCCGTCATCTATCCGCACT 697
 QY 173 ProThrGlyAsn---IleSerArgArgAlaProPheProTyrValAsnHisSerProAsn 191
 DB 698 CCAACATCGAGCTATGTTGATCCCAAGCGTTCACGATTCCTGTG-----AGGATTGTGGAC 751
 QY 192 ProSerArgGluPhe---SerGlySerIleGlyAsnValAlaIleTrpLysGluArgValAsp 210
 DB 752 CCCTCGAGAGACTTGAATTCCTATGGGCTTAATAGTGTGATCGAAGGAAAGAGTTGAG 811
 QY 211 GlyTrpLysMetLysGlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAla 230
 DB 812 AGCTGGAGGGTTAAACAGGACAAATAATATGTTGCAAGTGAATA-----856
 QY 231 ProSerGluGlyArgAlaIleThrAspIleAspAlaSerThrGluTyrAsnMetGluAsp 250
 DB 857 ---AAATATCCAGAGGCTAGAGGAGACATGGAG---GGGACTGGCTCAAATGGAGAGAT 910

QY 251 AlaLeuLeuAsnAspGluThrArgGlnProLeuSerArgLysValProIleAlaSerSer 270
 DB 911 ANGCAAAATGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 970
 QY 271 LysIleAsnProTyrArgMetValIleValLeuArgLeuValValLeuSerIlePheLeu 290
 DB 971 CAGCTCAACCTTTACCGGATAGTAATCATCTCTCTTATCATCTCTCTCTCTCTCTCTCT 1030
 QY 291 HisTyrArgLeuThrAsnProValArgAsnAlaIleTyrProLeuTrpLeuLeuSerValIle 310
 DB 1031 CAATATCGTATCAGTCATCAGTCATCAGTCATCAGTCATCAGTCATCAGTCATCAGTCATC 1090
 QY 311 CysGluIleTrpPheAlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIle 330
 DB 1091 TGTGAGGTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1150
 QY 331 AsnArgGluThrTyrLeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSer 350
 DB 1151 AACCGTGGAGACATATCTCGACAGGCTTGCAATGAGGTATGATAGAGAGGAGGCCATCA 1210
 QY 351 GlnLeuAlaAlaValAspIlePheValSerThrValAspProLeuLysGluProIle 370
 DB 1211 CAGCTGGCTCCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1270
 QY 371 ValThrAlaAsnThrValLeuSerIleLeuAlaValAspTyrProValAspLysValSer 390
 DB 1271 ATCAGAGCAACACACTGTTTGTCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1330
 QY 391 CysTyrValSerAspGlyAlaSerMetMetLeuThrPheAspAlaLeuAlaGluThrSer 410
 DB 1331 TGCTATGTTTCTGATGATGGCTCAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1390
 QY 411 GluPheAlaArgLysTrpValProPheValLysLysTyrAspIleGluProArgAlaPro 430
 DB 1391 GAATTTGCTAGAAAGTGGTTCCTTTTGAAGAAGCACAAATATTAAGAACCAAGAGCTCA 1450
 QY 431 GluPheTyrPheCysGlnLysIleAspTyrLeuLysAspLysValGlnProSerPheVal 450
 DB 1451 GAAATTTACTTTTGTCTCAAAAAATAGATTACCTGGAAGGACAAAATTCACCTTCATTTGT 1510
 QY 451 LysAspArgArgAlaMetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeu 470
 DB 1511 AAGGAAGACGAGCAATGAAGAGAGATGAAGAATTCAAAATAAGAATCAATGCCCTT 1570
 QY 471 ValSerLysAlaLeuLysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrp 490
 DB 1571 GTTGCCAAAGCACAGAAAGTCCCTGAGAGGGGTGGACCATGGCTGTTCTTGGGGCACAGTGT 1630
 QY 491 ProGlyAsnAsnThrArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGly 510
 DB 1631 CCTGGGAATAACCTTAGGGACCATCTGGCATGATTCAAGGTGTTCTTGGGGCACAGTGT 1690
 QY 511 GlyLeuAspThrGluGlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArg 530
 DB 1691 GGGCTTGACCTGATGAAATGAATTAACCACTCTTCTGCTGCTCTCTCTCTCTCTCTCTCT 1750
 QY 531 ProGlyPheGlnHisHisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAla 550
 DB 1751 CCAGCTTTCAGCATCAAGAGAGCTGGTGCAATGAATGACCTGATGCTGCTGCTGCTGCTG 1810
 QY 551 ValLeuThrAsnGlyGlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSer 570
 DB 1811 GTGCTGACAAATGGTGCCTATCTCTCAATGTGATGTGATGTGATGTGATGTGATGTGATGT 1870
 QY 571 LysAlaValArgGluAlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCys 590
 DB 1871 AAAGCTCTTAGAAGCAATGTGCTTTCATGATGATCCAGCTCTAGGAAGGAAACTTGT 1930
 QY 591 TyrValGlnPheProGlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArg 610
 DB 1931 TATGTACATTTCCACAAAGATTGTGATGCTTGCACGATGATGCTGCTGCTGCTGCTGCTG 1990

```

QY 611 AsnThrValPhePheAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyr 630
DB 1991 AACATAGTCTCTTTGATATCAACATGAAGAGCTTAGATGGATTCAGGTCACATCTAT 2050
QY 631 ValGlyThrGlyCysValPheAsnArgThrAlaIleTyrGlyTyrGluProPheIleLys 650
DB 2051 GTGGGAAACAGGATGCTGTTTCAATAGGACGGCTTTGTATGGATATGATCCTCTTTGACT 2110
QY 651 AlaLys-----LysProGlyPheLeuAlaSerLeuCysGlyGlyIleLysLysAlaSer 668
DB 2111 GAAGCTGATCTGGAACCTAACATCTGTTTAAAGAGCTGCTGGTAGAGGAAG---AGA 2167
QY 669 LysSerLysLysArgSerSerAspLysLysLysSerAsnLysHisValAspSerSerVal 688
DB 2168 AAGAACAAGAGTATATGATGATAGTCAAGCCGATATATCAAGACACAGAACTTCAGCT 2227
QY 689 ProValPheAsnLeuGluAspIleGluGluGlyValGluGlyAlaGlyPheAspAspGlu 708
DB 2228 CCATCTTTTAAATGAAGACATCGAGGAGGTATTGAA-----GGTTATGAGGATGAA 2281
QY 709 LysSerValLeuMetSerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPhe 728
DB 2282 AGGTCAAGTCTTATGTCCTCCAGAGAAATTTGAGAAACGCTTTGGTCAGTCTCCATCTTC 2341
QY 729 ValAlaSerThrLeuMetGluTyrGlyValProGlnSerSerThrProGluSerLeu 748
DB 2342 ATTGCATCCACCTTTATGACTCAAGGTGGATACACCTTCAACAAACCAGCTTCTCTA 2401
QY 749 LeuLysGluAlaIleHisValIleSerCysGlyTyrGluAspLysSerGluThrGlyThr 768
DB 2402 CTGAAGGAGCTATCCATGTTATCAGCTGTGGGTAGCAGACAAACTGAATCGGGAAAA 2461
QY 769 GluIleGlyTrpIleTyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHis 788
DB 2462 GAGATTGGCTGGATCTATGTTTACAGTGTGGGTACAGAGGATATCTGACTGGGTTTAAATGCAT 2521
QY 789 AlaArgGlyTrpArgSerValTyrCysMetProLysArgProAlaPheLysGlySerAla 808
DB 2522 GCAAGAGCTGGCAATCAATCTACTGATGCCACCGACCTTGTTTCAAGGGTCTTGCA 2581
QY 809 ProIleAsnLeuSerAspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGlu 828
DB 2582 CCATCAATCTTTCTGATGCTCTTAAATCAGGTGCTCGTGGGTCTCTGGTCAGTGAA 2641
QY 829 IleLeuPheSerArgHisCysProLeuThrTyrGlyTyrGlyArgLeuLysPheLeu 848
DB 2642 ATTCTGTTAGCAGACATTCCTATATGTTATGTTATGCTTACATGCGGATTTGAAGCTTTG 2701
QY 849 GluArgPheAlaTyrIleAsnThrThrIleTyrProLeuThrSerLeuProLeuVal 868
DB 2702 GAGAGGCTGGCTTACATTAACACCATTTGTTATCCATCATCTGTTCCGCTTATCGCC 2761
QY 869 TyrCysIleLeuProAlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSer 888
DB 2762 TATTGTGCTCTCTGCTATCTGCTCTTACCAATAAATTTATCATCTCTGAGATTAGT 2821
QY 889 AsnLeuAlaSerIleThrPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeu 908
DB 2822 AATTATGCTGAATGTTCTCTCTCTTCTTACCAATAAATTTATCATCTCTGAGATTAGT 2881
QY 909 GluMetArgTrpSerGlyValGlyIleAspGluTrpTrpArgAsnGlnPheTrpVal 928
DB 2882 GAGCTCAGATGAGTGTGTTGGCATTTGAAGATTGTTGGAGAAATGACAGCTTTTGGGTT 2941
QY 929 IleGlyGlyIleSerAlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAla 948
DB 2942 ATTGTGGCACCTCTGCCATCTCTTTCGGGGTGTCCAGGGTCTGCTGAAAGTGTGGCT 3001
QY 949 GlyIleAspThrAsnPheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAla 968
DB 3002 GGGATTGATACCACTTACAGTTACCTCAAGGATCTGATGAGATGGGACATTGCT 3061
QY 969 GluLeuTyrMetPheLysTrpThrThrLeuLeuIleProProThrThrIleLeuIleIle 988

```

```

DB 3062 GAGCTATATGTTTCAAGTGGACCAAGTTTGTCTCACTCCCTCCGACCACTGTTCTTCTCATT 3121
QY 989 AsnMetValGlyValValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrp 1008
DB 3122 AACCTGGTGGATGTTGGCAGGAATTCGTATGCATTAACACAGCGGTACCAATCTCTGG 3181
QY 1009 GlyProLeuPheGlyLysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPhe 1028
DB 3182 GGTCCGCTCTTTGGAAAGCTGTTCTTCGATCTGGGTGATCCTCATCTCTACCCCTTC 3241
QY 1029 LeuLysGlyLeuMetGlyArgGlnAsnArgThrProThrIleValIleValTrpAlaVal 1048
DB 3242 CTCACAGGCTCTCATGGCAGGAGAACCCGACCAACCAATCTGTATCTGTTGGTCCATC 3301
QY 1049 LeuLeuAlaSerIlePheSerLeuLeuTrpValArgValAspProPhe-----Thr 1065
DB 3302 CTCCTTGGCTCTATCTTCTCTTGTGGGTGAAGATCGATCTTTCATCTCTCCCGACA 3361
QY 1066 ThrArgLeuAlaGlyProAsnIleGlnThrCysGlyIleAsnCys 1080
DB 3362 CAGAAAGCTGCCGCC-----TTGGGGCAATGTGGTGAAGTGC 3400
RESULT 13
AAZ99530
ID AAZ99530 standard; DNA; 3753 BP.
AC AAZ99530;
XX
XX 03-JUL-2000 (first entry)
DT
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
XX Key Location/Qualifiers
XX CDS 184..3405
XX /*tag= a
XX /product= "cellulose synthase"
XX /transl_except= (pos: 916..918, aa: Xaa)
XX /note= "no termination codon given; Xaa is an unspecified
XX amino acid"
XX
XX WC200009706-A2.
XX
XX 24-FEB-2000.
XX
XX 16-AUG-1999; 99WO-US018760.
XX
XX 17-AUG-1998; 98US-0096822P.
XX
XX (PTON-) PIONEER HI-BRED INT INC.
XX
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
XX WPI: 2000-224343/19.
XX
XX P-FSDB; AAY84120.
XX
XX New genes which encode maize cellulose synthase polypeptides in plants
XX useful for modulating the expression of cellulose synthase in plants and
XX to produce transgenic plants expressing the novel protein.
XX
XX Claim 1; Page 184-188; 119pp; English.
XX
XX The present sequence encodes a maize cellulose synthase polypeptide. The
XX cellulose synthase can be used for the improvement of stalk quality for
XX improved stand or silage. It also provides an increased concentration of
XX cellulose in the pericarp, hardening the kernel and improving its
XX handling ability. The sequences are used to produce transgenic plants and
XX seeds expressing the cellulose synthase. The polynucleotide is used for

```

modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein

SQ Sequence 3753 BP; 944 A; 805 C; 965 G; 1038 T; 0 U; 1 Other;

Alignment Scores:		
Pred. No.:	0	3753
Score:	3981.00	747
Percent Similarity:	80.64%	132
Best Local Similarity:	68.53%	163
Query Match:	68.90%	132
DB:	3	19
		Gaps: 19
		Indels: 48
		Mismatches: 132
		Conservative: 163
		Matches: 3753

US-09-900-237A-30 (1-1080) x AAZ99530 (1-3753)

2 AspGlyAspAla---AspAlaLeuIysSerGlyArgHisGlyAlaGlyAspValCysGln 20
Db GACGGCGATGTGCCGGGCTCGGCTAAGCCACAAAGAGTGCGAATGACAGGTCTGCCAG 309
21 IleCysAlaAspGlyLeuGlyThrThrLeuaspGlyAspValpHeThrAlaCysaspVal 40
Db ATTTGCGGTGACTCTGTGGGTGTTTCAGCCACTGGTGATGCTCTTTGTTGCCATGAG 369
41 CysArgpHeProValCysArgProCysTyrGluHisGluArgLysGluGlyThrGlnAla 60
Db TGTGCTTCCCTGTCTGCGGCCATGCTATGAGTATGAGCGCAAGGAGGGGAACCAATGC 429
61 CysLeuGlnCysLysThrLysTyrIysArgHisArgGlySerProAlaIleArgGlyGlu 80
Db TGCCCCAGTCCAAGACTAGATACAAAGACAGCAAGGATGACCCCTCGAGTTCATGTTGAT 489
81 GluGlyAspAspThrAspAlaAspAsp---GlySerAspPheAsnTyrProAlaSerGly 99
Db GAG---GATGAGGAAGATGTTGATGACCTAGACAAATGAATTCACATAC---534
100 ThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGly 119
Db 535 -----AGCAA-----540
120 GlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyr 139
Db 541 GGCAGTGGGAAAGGCCAGAGTGCCAACTGCAAGAGAGATGATGCTGCTCTCATCT 600
140 AspSerGlyGluIleProArgGlyTyrValProSerValThrAsn---SerGlnMetSer 158
Db 601 GCTCGCATAG---CCACATATCGGATTCACGCGCTGACGCGTGCACAGGTATCT 657
159 GlyGluIleProGlyAlaSerProAspHisHis---MetMetSerProThrGlyAsn---176
Db 658 GGAGAGATTCCTGATGCTTCCCTGACCGGTATCTATCCGAGTCCACATCGAGCTAT 717
177 IleSerArgArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe 196
Db 718 GTTGATCCAAAGCTGCCAGTTCCTGTG-----AGGATGTGACACCCTCGAAGGACTTG 771
197 ---SerGlySerIleGlyAsnValAlaTTrpLysGluArgValAspGlyTTrpLysMetLys 215
Db 772 AATTCTCTATGGGCTTAATAGTGTGTGACTGGAGGAAAGAGTTGAGACTGGAGGTTTAAA 831
216 GlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg 235

1894 GCAATGCTCTCATGATGGATCCGGCTCTAGGAAGAAAACCTTGTTATGTACTCAATTTC 1953
596 GlnArgPheAspGlyValAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe 615
1954 CAGAGATTGATGGCATTGACCTTGACAGATCGATATGCTTAATCGGAACATAGCTTTCTTT 2013
616 AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys 635
2014 GATATCAACATGAAAGGCTGATGATGCAATTCAGGGTCCAGTTTACGTGGGAACAGATGC 2073
636 ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleIleAlaLys-----Lys 653
2074 TGTTTCAATAGACAGCTTTGATGATAGATAGATCCCTGTTTGTACTGAAGCTGATCTGGAG 2133
654 ProGlyPheLeuAlaSerLeuCysGlyGlyLysLysLysLysLysLysLysLysLysLys 673
2134 CCAACATGTTATTAAAGAGCTGCTGTGTAGAGGAG--AAAAAGAACAAAGATTAT 2190
674 SerSerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu 693
2191 ATGATAGTCAAAAGCCGATTATGAAAGAGACAGAAATCTTCAGCTCCCATCTTCAATATG 2250
694 GluAspIleGluGlyValGluGlyAlaGlyPheAspAspGlyLysSerValLeuMet 713
2251 GAAGACATCGAAGAGGATTTGAA-----GGTTACGAGATGAAGGTCAGTGTCTATG 2304
714 SerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeu 733
2305 TCCAGAGGAAATGGAGAAACGCTTGGTCAGTCTCTATTATTTCAATGATCCACCTTT 2364
734 MetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLysGluAlaIle 753
2365 ATGACACAAGGTGGCATACCACTTCAACAAACCCAGCTTCTTACCTAAAGGAGCTATC 2424
754 HisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIle 773
2425 CATGTCATCAGTTGTGATATCAGACAAACTGAATGGGAAAGAGATTGGCTGATC 2484
774 TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArg 793
2485 TATGTTTCAGTAACGGAGGATATTCTGACTGGTTTAAATGTCATGCAAGGGCTGGCAA 2544
794 SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer 813
2545 TCAATCTACTGATGCCACCAACGCTGTTTCAAGGGTTCTGCACCAATCAATCTTCC 2604
814 AspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArg 833
2605 GATGCTCTTAATCAGTGTCTCGTTGGCTCTTGGGTGAGTGAATTCGTTAGTAGA 2664
834 HisCysProLeuTyrTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyr 853
2665 CATTTGCTCTATCTGTTGTTTACAATGACGATTAAGAGCTTTTGGAGGGCTGGCTTAC 2724
854 IleAsnThrThrIleTyrProLeuThrSerLeuProLeuValTyrCysIleLeuPro 873
2725 ATCAACACTATTGTATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2784
874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle 893
2785 GCTATCTGCTCTCTTACCAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2844
894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSer 913
2845 TTCTTCACTCTTCTTTGCGCTCCATTTTGGCACTGGTATATGAGCTTAGATGAGAT 2904
914 GlyValGlyIleAspGluTrpTrpArgAsnGluGlnPheThrValIleGlyIleSer 933
2905 GGTGTGGGATGAAAGATTGGTGGAAATGAGCAGTTTGGGTATTGGTGGCACCCTCT 2964
934 AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn 953
2965 GCCCATCTCTTCGACGATGTTCCAGGGTCTGCTGAAAGTGTGGCTGGGATTGATACCAAC 3024
954 PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe 973
3025 TTCACAGTTACCTCAAGGCATCTGATGAGGATGGCGACTTTGCTGAGCTATATGTGTTTC 3084
974 LysTrpThrThrLeuLeuIleProProThrThrIleLeuIleIleAsnMetValGlyVal 993
3085 AAGTGGACCCAGTTTGTCTATCTCCGACCATCTGTTTGTTCATTAACCTGGTCCGAATG 3144
994 ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGly 1013
3145 GTGGCAGGAAATTTCTTATGCCATTAACAGTGGCTACCAATCTCTGGGGTCCGCTCTTTGA 3204
1014 LysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMet 1033
3205 AAGCTGTTCTTCTCGATCTGGGTGATCTCTCCATCTCTACCCCTTCTCAAGGGTCTCATG 3264
1034 GlyArgGlnAsnArgThrProThrIleValIleValIleValIleValIleValIleVal 1053
3265 GGAAAGCAGAACCGCACCAACAATCGTCATTTGTTCTGTTCCATCTCTTGTGATCTATC 3324
1054 PheSerLeuLeuTyrValArgValAspProPhe-----ThrThrArgLeuAlaGly 1070
3325 TTCTCTTGTGTTGGGTGAAGATCGATCTTTTCATCTCCCGCACACAGAAAGCTGCTGCC 3384
1071 ProAsnIleGlnThrCysGlyIleAsnCys 1080
3385 -----TTGGGCAATGTGGCGTCAACTGC 3408
RESULT 14
AAZ99515
ID AAZ99515 standard; DNA; 3753 BP.
AC AAZ99515;
XX
XX 03-JUL-2000 (first entry)
DE DNA encoding a maize cellulose synthase.
XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX Zea mays.
OS
XX
XX Key Location/Qualifiers
XX CDS 184..3405
XX /*tag= a
XX /product= "cellulose synthase"
XX /transl_except=(pos: 916..918, aa: Xaa)
XX /note= "no termination codon given; Xaa is an unspecified
XX amino acid"
XX
XX WO200009706-A2.
XX
XX 24-FEB-2000.
XX
XX 16-AUG-1999; 99WO-US018760.
XX
XX 17-AUG-1999; 98US-0096822P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
XX WPI: 2000-224343/19.
XX P-PSDB; AAY84115.
XX
XX New genes which encode maize cellulose synthase polypeptides in plants
XX useful for modulating the expression of cellulose synthase in plants and
XX to produce transgenic plants expressing the novel protein.
XX
XX Claim 1; Page 144-149; 119pp; English.
XX

CC The present sequence encodes a maize cellulose synthase polypeptide. The
 CC cellulose synthase can be used for the improvement of stalk quality for
 CC improved stand or silage. It also provides an increased concentration of
 CC cellulose in the pericarp, hardening the kernel and improving its
 CC handling ability. The sequences are used to produce transgenic plants and
 CC seeds expressing the cellulose synthase. The polynucleotide is used for
 CC modulating, preferably increasing, the level of the synthase in a plant
 CC cell. The plants are preferably monocots. The polynucleotide is also used
 CC as a probe or primer in the detection quantitation or isolation of gene
 CC transcripts. The probes are useful in detecting deficiencies in the level
 CC of mRNA in screenings for desired transgenic plant, for detecting
 CC mutations in the gene, for monitoring upregulation of expression or
 CC changes in enzyme activity in screening assays of compounds, for
 CC detection of any number of allelic variants of the gene, or for use as
 CC molecular markers in plant breeding programs. The isolated nucleic acids
 CC of the present invention can also be used for recombinant expression of
 CC their encoded polypeptides or for use as immunogens in the preparation
 CC and/or screening of antibodies. The proteins can be employed in assays
 CC for enzyme agonists or antagonists of enzyme function or for use of
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with a protein
 CC
 XX

SQ Sequence 3753 BP; 944 A; 805 C; 965 G; 1038 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 0 Length: 3753
 Score: 3981.00 Matches: 747
 Percent Similarity: 80.64% Conservative: 132
 Best Local Similarity: 68.53% Mismatches: 163
 Query Match: 68.90% Indels: 48
 DB: 3 Gaps: 19

US-09-900-237A-30 (1-1080) x AAZ99515 (1-3753)

CC	197	---	SerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTyrLysMetLys	215
CC	772	AATTCCTATGGGCTTAATAGTGTGACTGGAAGAAAGAGTGTGAGAGCTGGAGGGTTAAA	831	
CC	216	GlnAspLysGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg	235	
CC	832	CAGACAAAAAATATGATGCAAGTGACTAAT-----AAATATCCAGAGGCTAGA	879	
CC	236	AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp	255	
CC	880	GGAGGA---GACATGGAG---GGGACTGGCTCAATGGAGAAATATGCAATGGTGGAT	933	
CC	256	GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr	275	
CC	934	GATGCACGGCTACCTTTGAGCGGTATCGTGCCAATTTCTCTCAACAGCTCAACCTTTAC	993	
CC	276	ArgMetValIleValLeuArgLeuValLeuSerIlePheLeuHisTyrArgLeuThr	295	
CC	994	CGGTAGTGTATCATCTCCGTCTTATCATCCTGTGCTTCTTCTTCTCAGTATCGTGTAGT	1053	
CC	296	AsnProValArgAsnAlaTyrProLeuTyrProLeuLeuSerValIleCysGluIleTyrPhe	315	
CC	1054	CATCCAGTGGTGATGCTTATGGATTTATGGTAGTATCTGTATCTGCCAGGCTCGTTT	1113	
CC	316	AlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyr	335	
CC	1114	GCCTTGTCTGGCTTCTAGATCAGTTCCTCCAAATGGTATCCAATCAACCGTGAGACATAT	1173	
CC	336	LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaVal	355	
CC	1174	CTTGACAGGCTTGCAATTGAGGTATGATAGAGGGAGAGCCATCAGCTGGTCCCAT	1233	
CC	356	AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr	375	
CC	1234	GATGCTTCCTCGTACGTACGTGGATCCATTAAGAGAACCTTCCACTGATCAGCCACACT	1293	
CC	376	ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp	395	
CC	1294	GTTTTGTCCATCTTCTTGTGGATTACCTGTGACAAAGTGTCTATGCTATGTTCTGAT	1353	
CC	396	AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys	415	
CC	1354	GATGTTGAGCTATGCTGACTTTTTCAGTCTCTCTCAGAAACCGCAGAAATTTCTTGAAG	1413	
CC	416	TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435	
CC	1414	TGGGTTCCCTTTTGTAAAGAGCACAAATTTGACCCAGAGCTCCAGAAATTTTACTTTGCT	1473	
CC	436	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455	
CC	1474	CAAAAAATAGATTACTGAAGACAAAAATCAACCTTCATTTGTTAAGAAAGACGCGCA	1533	
CC	456	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475	
CC	1534	ATGAAGAGGAGTAGAAGAAATCAAGTAAGATCAATGCCCTTGTTCGCAAGACACAG	1593	
CC	476	LysValProGluGluGlyTyrIleMetGlnAspGlyThrProTyrProGlyAsnAsnThr	495	
CC	1594	AAAGTCTCGAAGGGGGTGACCATGGCTGATGGAACTGCATGGCTCGGGGAATAATCCT	1653	
CC	496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515	
CC	1654	AGGACCATCTCGCATGATTCAGGTTTCTTGGGGCACAGTGGTGGCTCGACACTGAT	1713	
CC	516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535	
CC	1714	GGAAATGAGTTACACCGTCTGTCTATGTCTCTCTCGTGTAAAAAGACCGAGCTTTT	1773	
CC	536	HisLysLysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly	555	
CC	1774	CACAAGAGGCTGTGTCAATGAATGCCCTGATTCGTATCTGTCTGTGCTGACAAATGGT	1833	

QY	556	GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu	575	Db	2905	GGTGTGGCATTGAAGATTGGTGGAGAAATGAGCAGTTTGGGTATTGGTGGACCTCT	2964
Db	1834	GCCTATCTCTCAATGTGATGGCGACCAATTAATCAATAGCAGCAAGCTCTTAGAGAA	1893	QY	934	AlaHisLeuPheAlaValPheGlnGlyLeuLeuLysValLeuAlaGlyIleAspThrAsn	953
QY	576	AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro	595	Db	2965	GCCCATCTCTTCGACGTTCGAGGGTCTGCTGAAAGTGTGGCTGGGATTGATCAAC	3024
Db	1894	GCAATGTGCTTCATGATGATCGGCTCTAGGAGGAAACCTTGTATGTACAAATTC	1953	QY	954	PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe	973
QY	596	GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe	615	Db	3025	TTACAGATTACCTCAAGGCACTGATGAGGATGGCGACTTGTGCTGAGCTATATGTGTT	3084
Db	1954	CAGAGATTTGATGGCATTGACCTTGACGATCGATATGCTAATCGGAACATAGTTTCTTT	2013	QY	974	LysTyrThrThrLeuLeuLeuProProThrThrIleLeuIleLeuMetValGlyVal	993
QY	616	AspIleAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys	635	Db	3085	AAGTGGACCAAGTTTGTCTCATTCCTCGACACCTGTTCTTGTCAATTAACCTGGTGGAA	3144
Db	2014	GATATCAACATGAAGGCTCTGGATGGCATTACGGGTCCAGTTTACGTGGGAACAGGATC	2073	QY	994	ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTyrGlyProLeuPheGly	1013
QY	636	ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLys-----Lys	653	Db	3145	GTGGCAGGAATTTCTTATGCCATTAAACAGTGGCTACCAATCCTGGGGTCCGCTTTTGA	3204
Db	2074	TGTTTCAATAGACAGGCTTGTATGATACGATCCTGTTTGTACTGAAGCTGATCTGGAG	2133	QY	1014	LysLeuPhePheAlaPheTyrValIleValHisLeuTyrProPheLeuLysGlyLeuMet	1033
QY	654	ProGlyPheLeuAlaSerLeuCysGlyGlyLysLysAlaSerLysSerLysLysArg	673	Db	3205	AAGCTGTTCTTCGATCTGGGTGATCCTCACTCTACCCCTTCCTCAAGGGTCTCATG	3264
Db	2134	CCAAACATTTGTTAAGAGCTGCTGTGTAGAGGAAG---AAAAGAAACAAGAGTTAT	2190	QY	1034	GlyArgGlnAsnArgThrProThrIleValIleValTyrAlaValLeuLeuAlaSerIle	1053
QY	674	SerSerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu	693	Db	3265	GGAAGGCAGAACCGCACCAACAATCGTCAATGTCTGTGGTCCATCTCTTGTGATCTATC	3324
Db	2191	ATGGATAGTCAAAAGCCGTTATATGAAGAGAACAGAACTCTCAGCTCCATCTTCAATG	2250	QY	1054	PheSerLeuLeuTrpValArgValAspProPhe-----ThrThrArgLeuAlaGly	1070
QY	694	GluAspIleGluGluGlyValGluGlyAlaGlyPheAspAspGluLysSerValLeuMet	713	Db	3325	TTCTCTTGTCTGTGGGTGAAGATCGATCTTTTATCTCCCCGACACAGAAAGCTGCTGCC	3384
Db	2251	GAAGACATCGAAGAGGTTATGAA-----GGTTACGAGGATGAAGGTCAGTCTTATG	2304	QY	1071	ProAsnIleGlnThrCysGlyIleAsnCys	1080
QY	714	SerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaPheValAlaSerThrLeu	733	Db	3385	-----TTGGGGCAATGTGGCGTCAACTGC	3408
Db	2305	TCCGAGAGGAATTTGGAGAAACGCTTGTGTCAGTCTCTCTATTTTTCATTGCTCACCTTT	2364	RESULT 15			
QY	734	MetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIle	753	AAZ99497			
Db	2365	ATGACACAAAGTGGCATACCACTTCAACAAACCCAGCTTCTCTACTTAAGGAAGCTATC	2424	ID	AAZ99497	standard; DNA; 3780 BP.	
QY	754	HisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIle	773	XX	AAZ99497;		
Db	2425	CATGTCATCAGTTGTGATATGAGGACAAACTGAATGGGGAAGAGATTGGCTGGATC	2484	AC			
QY	774	TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArg	793	DT	03-JUL-2000	(first entry)	
Db	2485	TATGGTTCAGTAACGAGGATATCTGACTGGGTTTAAATGCAATGCAAGCGGCTGGCAA	2544	DE		DNA encoding a maize cellulose synthase.	
QY	794	SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer	813	KW		Maize; cellulose synthase; stalk quality; stand; silage; cellulose;	
Db	2545	TCAATCTACTGATGCCACCGACCTTGTTCAGGGTTCTGACCCATCAATCTTTCC	2604	KW		transgenic plant; plant breeding marker; ss.	
QY	814	AspArgLeuAsnGlnValLeuArgTrpAlaLeuGlySerValGluIleLeuPheSerArg	833	XX		Zea mays.	
Db	2605	GATCGTCTTATCAGTGTCTCGTGGGCTCTTGGTCAAGTGGAAATCTGCTTAGTAGA	2664	XX		Location/Qualifiers	
QY	834	HisCysProLeuTyrTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyr	853	XX		201..3422	
Db	2665	CATTGTCTCTATCTGTTATGTTTACAAATGAGGATTTGAAGCTTTTGGAGGGCTGCTTAC	2724	XX		/*tag=	
QY	854	IleAsnThrThrIleTyrProLeuThrSerLeuProLeuLeuValTyrCysIleLeuPro	873	XX		/product= "cellulose synthase"	
Db	2725	ATCAACATATTGTATATCCATCAATCCATCCGCTTATGCTTATGTGCTTCC	2784	XX		/trans= except= (pos: 933..935, aa: Xaa)	
QY	874	AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle	893	XX		/note= "no termination codon given; Xaa is an unspecified	
Db	2785	GCTATCTGCTCTTACCAATAAATTTATCATCTCTGAGATTAGCAATTTATCTGGATG	2844	XX		amino acid"	
QY	894	TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTrpSer	913	XX		WO200009706-A2.	
Db	2845	TTCTTCAATCTCTTTTCCGCTCCATTTTGGCATGGTATATTTGAGCTTAGATGGAGT	2904	XX		24-FEB-2000.	
QY	914	GlyValGlyIleAspGluTrpTrpArgAsnGluGlnPheTrpValIleGlyIleSer	933	XX		16-AUG-1999; 99WO-US018760.	
				XX		17-AUG-1998; 98US-0096822P.	
				XX		(PION-) PIONEER HI-BRED INT INC.	
				XX		Dhugga KS, Helentjaris TG, Bowen BA, Wang X;	
				XX		WPI; 2000-224343/19.	
				XX		P-PSDB; AAY84109.	

PT New genes which encode maize cellulose synthase polypeptides in plants
 PT useful for modulating the expression of cellulose synthase in plants and
 PT to produce transgenic plants expressing the novel protein.

XX Claim 1; Page 97-102; 119pp; English.

XX The present sequence encodes a maize cellulose synthase polypeptide. The
 CC cellulose synthase can be used for the improvement of stalk quality for
 CC improved stand or silage. It also provides an increased concentration of
 CC cellulose in the pericarp, hardening the kernel and improving its
 CC handling ability. The sequences are used to produce transgenic plants and
 CC seeds expressing the cellulose synthase. The polynucleotide is used for
 CC modulating, preferably increasing, the level of the synthase in a plant
 CC cell. The plants are preferably monocots. The polynucleotide is also used
 CC as a probe or primer in the detection quantitation or isolation of gene
 CC transcripts. The probes are useful in detecting deficiencies in the level
 CC of mRNA in screenings for desired transgenic plant, for detecting
 CC mutations in the gene, for monitoring upregulation of expression or
 CC changes in enzyme activity in screening assays of compounds, for
 CC detection of any number of allelic variants of the gene, or for use as
 CC molecular markers in plant breeding programs. The isolated nucleic acids
 CC of the present invention can also be used for recombinant expression of
 CC their encoded polypeptides or for use as immunogens in the preparation
 CC and/or screening of antibodies. The proteins can be employed in assays
 CC for enzyme agonists or antagonists of enzyme function or for use of
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with a protein

SQ Sequence 3780 BP; 946 A; 817 C; 976 G; 1040 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	0	Length:	3780
Score:	3981.00	Matches:	747
Percent Similarity:	80.64%	Conservative:	132
Best Local Similarity:	68.53%	Mismatches:	163
Query Match:	68.90%	Indels:	48
DB:	3	Gaps:	19

US-09-900-237A-30 (1-1080) x AA299497 (1-3780)

QY	2	AspGlyAspAla--AspAlaLeuLysSerGlyArgHisGlyAlaGlyAspValCysGln	20
DB	267	GACGGGATGTCGGGGTCGGCTAAGCCCAAGAGAGTGCGAATGGACAGGTCTGCCAG	326
QY	21	IleCysAlaAspGlyLeuGlyThrThrLeuAspGlyAspValPheThrAlaCysAspVal	40
DB	327	ATTTCGGGTGACTCTGTGGGTGTTTCAGCCACTGGTGATGCTTTGTTCCTGCAATGAG	386
QY	41	CysArgPheProValCysArgProCysTyroGluHisGluArgLysGluGlyThrGlnAla	60
DB	387	TGTGCTTCCTGCTGCTCCGCCCATGCTATGAGTATGAGCAAGGAGGGGAACCAATGC	446
QY	61	CysLeuGlnCysLysThrLysTyroLysArgHisArgGlySerProAlaIleArgGlyGlu	80
DB	447	TGCCCCAGTGCACAGATAGATACAGAGACAGAAAGGTAGCCCTCGAGTTCATGTTGAT	506
QY	81	GluGlyAspThrAspAlaAspAsp--GlySerAspPheAsnTyrProAlaSerGly	99
DB	507	GAG--GATGAGGAAGATGTTGATGACCTAGACAAATGAATCAACTAC-----	551
QY	100	ThrGluAspGlnLysGlnLysIleAlaAspArgMetArgSerTrpArgMetAsnThrGly	119
DB	552	-----AAGCAA-----	557
QY	120	GlySerGlyAsnValGlyHisProLysTyrAspSerGlyGluIleGlyLeuSerLysTyr	139
DB	558	GGCAGTGGGAAGGCCAGAGTGGCAACTGCACAGGAGATGATGCTGCTCTCTCTCTCT	617
QY	140	AspSerGlyGluIleProArgGlyTyrValProSerValThrAsn--SerGlnMetSer	158
DB	618	GCTCGCCATGAG--CCACATCATCGGATTCACGGCTGACAGCGGTCAACAGATATCT	674
QY	159	GlyGluIleProGlyAlaSerProAspHisHis--MetMetSerProThrGlyAsn---	176

DB	675	GGAGAGATTCTGATGCTTCCCTGACCGTCATTTCTATCGCAGTCCCAATCGAGCTAT	734
QY	177	IleSerArgArgAlaProPheProTyrValAsnHisSerProAsnProSerArgGluPhe	196
DB	735	GTTGATCCCAAGCGTCCAGCTTCTGTG-----AGGATTGTGACCCCTCGAAGGACTTG	788
QY	197	---SerGlySerIleGlyAsnValAlaTrpLysGluArgValAspGlyTrpLysMetLys	215
DB	789	AATTCTTATGGGCTTAATAGTGTGATGGAAGAAAGAGTTCAGAGCTGAGAGGTTAAA	848
QY	216	GlnAspGlyGlyAlaIleProMetThrAsnGlyThrSerIleAlaProSerGluGlyArg	235
DB	849	CAGGACAAATAATGATGCAAGTCACTAAT-----AAATATCCAGAGGCTAGA	896
QY	236	AlaAlaThrAspIleAspAlaSerThrGluTyrAsnMetGluAspAlaLeuLeuAsnAsp	255
DB	897	GGAGGA--GACATGGAG--GGGACTGGCTCAATGGAGAAATATGCAAAATGGTTGAT	950
QY	256	GluThrArgGlnProLeuSerArgLysValProIleAlaSerSerLysIleAsnProTyr	275
DB	951	GATGACGGCTACCTTTGAGCGGTATCGTGCCTAATTTCTCAACAGCTCAACCTTTAC	1010
QY	276	ArgMetValIleValLeuArgLeuValValLeuSerIlePheLeuHisTyrArgLeuThr	295
DB	1011	CGGGTAGTATCATTTCTCGTCTTATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1070
QY	296	AsnProValArgAsnAlaTyrProLeuTrpLeuLeuSerValIleCysGluIleTrpPhe	315
DB	1071	CATCCAGTGGGTGATGCTTATGGATTATGGCTAGTATCTGTTATCTGGAGGTCTGGTTT	1130
QY	316	AlaLeuSerTrpIleLeuAspGlnPheProLysTrpPheProIleAsnArgGluThrTyr	335
DB	1131	GCCTTGTCTGGCTTCTAGATCAGTTCCCAAAATGGTATCCCAATCAACGGTGAGACATAT	1190
QY	336	LeuAspArgLeuAlaLeuArgTyrAspArgGluGlyGluProSerGlnLeuAlaVal	355
DB	1191	CTTGACAGCTTGCATTGAGGTATGATAGAGGGAGAGCCATCACAGCTGGCTCCCAT	1250
QY	356	AspIlePheValSerThrValAspProLeuLysGluProProIleValThrAlaAsnThr	375
DB	1251	GATGCTTCGTGTCAGTACAGTGGATCCATTGAAGAACCTCCACTGATCACAGCCAACT	1310
QY	376	ValLeuSerIleLeuAlaValAspTyrProValAspLysValSerCysTyrValSerAsp	395
DB	1311	GTTTGTGTCATTCTTCTGTGATTACCTCTGTGACAAAGTGTCTGCTATGTTCTGAT	1370
QY	396	AspGlyAlaSerMetLeuThrPheAspAlaLeuAlaGluThrSerGluPheAlaArgLys	415
DB	1371	GATGGTTGAGCTATGCTGACTTTTGGAGTCTCTCTCAGAAACCGCAGAAATTTGCTAGA	430
QY	416	TrpValProPheValLysLysTyrAspIleGluProArgAlaProGluPheTyrPheCys	435
DB	1431	TGGGTTCCCTTTGTAGAGAGCAATATTGAACCAAGAGCTCCAGAAATTTTACTTTGCT	1490
QY	436	GlnLysIleAspTyrLeuLysAspLysValGlnProSerPheValLysAspArgAla	455
DB	1491	CAAAAAATAGATTACCTGGAAGGACAAATTCACCTTCATTGTTTAGGAAAGCGCGCA	1550
QY	456	MetLysArgGluTyrGluGluPheLysIleArgIleAsnAlaLeuValSerLysAlaLeu	475
DB	1551	ATGAAGGGGAGTATGAAGAATCAAGTAAGAAATCAATGCTCTGTTGTCGAAGCACAG	1610
QY	476	LysValProGluGluGlyTrpIleMetGlnAspGlyThrProTrpProGlyAsnAsnThr	495
DB	1611	AAAGTCCCTGAAGAGGGGTGGACCATGGCTGATGGAACTGCATGGCTGGGAATAATCT	1670
QY	496	ArgAspHisProGlyMetIleGlnValPheLeuGlyHisSerGlyGlyLeuAspThrGlu	515
DB	1671	AGGGACCATCTGGCATGATTGAGGTTTCTTGGGGCAGCAGTGGTGGGCTCGACATGAT	1730
QY	516	GlyAsnGluLeuProArgLeuValTyrValSerArgGluLysArgProGlyPheGlnHis	535

1731 GGAATGAGTTACACGCTTGTCTATGCTCTCTCGTGAAGAGACAGGCTTTCCAGCAT 1790
 536 HisLysIysAlaGlyAlaMetAsnAlaLeuValArgValSerAlaValLeuThrAsnGly 555
 1791 CACAGAGGCTGGTGCATGAATGCGCTGATTCGTGTATCTGCTGCTGACAAATGGT 1850
 556 GlnTyrMetLeuAsnLeuAspCysAspHisTyrIleAsnAsnSerLysAlaValArgGlu 575
 1851 GCCTATCTTCAATGTGGATTGCGACCATTAATCAATAGCAGAAAGCTCTTGAAGAA 1910
 576 AlaMetCysPheLeuMetAspProAsnLeuGlyProGlnValCysTyrValGlnPhePro 595
 1911 GCAATGTGCTTCAATGATGATCCGCTCTAGAGAGAAACTTGTATGTAACAATTTCCA 1970
 596 GlnArgPheAspGlyIleAspArgAsnAspArgTyrAlaAsnArgAsnThrValPhePhe 615
 1971 CAGAGATTGATGGCATTCACITTCGACGATGATATGCTTAATCGGAACATAGTTTCTTT 2030
 616 AspileAsnLeuArgGlyLeuAspGlyIleGlnGlyProValTyrValGlyThrGlyCys 635
 2031 GATATCAACATGAAGGCTGATGCGATTCAGGGTCCAGTTTACGTGGGAACAGATGC 2090
 636 ValPheAsnArgThrAlaIleTyrGlyTyrGluProProIleLysAlaLys-----Lys 653
 2091 TGTTCATATAGACAGGCTTTGTATGATACGATCTGTTTGTACTGAAGTGTATCGGAG 2150
 654 ProGlyPheLeuAlaSerLeuCysGlyGlyLysLysAlaSerLysSerLysLysArg 673
 2151 CCMAACATTGTTATTAGAGCTGCTGTGGTGAAGGAAG---AAAAAGAACAGAGATTAT 2207
 674 SerSerAspLysLysLysSerAsnLysHisValAspSerSerValProValPheAsnLeu 693
 2208 ATGGATAGTCAAAAGCCGTATTATGAAGAGAACAGAAATCTTCAGTCCCATCTTCAATATG 2267
 694 GluAspIleGluGluGlyValGluGlyAlaGlyPheAspAspGluLysSerValLeuMet 713
 2268 GAAGACATCGAAGAGGTTATGAA-----GGTTACGAGGATGAAGGTCAGTGTCTATG 2321
 714 SerGlnMetSerLeuGluLysArgPheGlyGlnSerAlaAlaPheValAlaSerThrLeu 733
 2322 TCCAGAGGAATGGAGAAACGCTTGTGTCAGTCTCTATTTTCATTTGATGCCACCTTT 2381
 734 MetGluTyrGlyGlyValProGlnSerSerThrProGluSerLeuLeuLysGluAlaIle 753
 2382 ATGACACAAGGTGGCATACCACTTCAACAAACCCAGCTTCTCTACTTAAGAGAGCTATC 2441
 754 HisValIleSerCysGlyTyrGluAspLysSerGluTyrGlyThrGluIleGlyTyrIle 773
 2442 CATGTCAATCAGTTGTGATATGAGACAAAACTGAATGGGGAAGAGATTGGCTGGATC 2501
 774 TyrGlySerValThrGluAspIleLeuThrGlyPheLysMetHisAlaArgGlyTyrArg 793
 2502 TATGTTTCAGTAACGAGGATATCTGACTGGGTTTAAATGCAATGCAAGGGGCTGGCAA 2561
 794 SerValTyrCysMetProLysArgProAlaPheLysGlySerAlaProIleAsnLeuSer 813
 2562 TCAATCTACTGTCATGCCACCAAGGCTTGTTCAGGGTTCTGACCAATCAAICTTTCC 2621
 814 AspArgLeuAsnGlnValLeuArgTyrAlaLeuGlySerValGluIleLeuPheSerArg 833
 2622 GATGTCCTTAATCAGGTGCTCCGTTGGGCTCTTGGGTGAGTGGAAATTCGTAGTAGA 2681
 834 HisCysProLeuTyrTyrGlyTyrGlyArgLeuLysPheLeuGluArgPheAlaTyr 853
 2682 CATTCGTCTATCTGATGTTTACAAATGACGATTAAGAGCTTTTGGAGAGGCTGGCTTAC 2741
 854 IleAsnThrIleTyrProLeuThrSerLeuLeuValTyrCysIleLeuPro 873
 2742 ATCAACACATTGTTATATCCAAATCAATCCATCCGCTTATGCTGCTGCTTCC 2801
 874 AlaIleCysLeuLeuThrGlyLysPheIleMetProGluIleSerAsnLeuAlaSerIle 893
 2802 GCTATCTGCTCTTACCAATAAATTTATCAATCTGAGATTAGCAATTTATGCTGGGATG 2861

894 TrpPheIleAlaLeuPheLeuSerIlePheAlaThrGlyIleLeuGluMetArgTyrSer 913
 2862 TTCCTTCACTTCTTTTCGCCTCCATTTTCCCACTGGTATATTGGAGCTTAGATGGAGT 2921
 914 GlyValGlyIleAspGluTyrTrpArgAsnGluGlnPheTrpValIleGlyGlyIleSer 933
 2922 GGTGTGGCATGTGAAGATTGGTGGAGAAATGAGCAGATTTTGGGTATTGGTGGCACCTCT 2981
 934 AlaHisLeuPheAlaValPheGlnGlyLeuLysValLeuAlaGlyIleAspThrAsn 953
 2982 GCGCATCTCTTCGCAAGTGTTCAGGGTCTGCTGAAAGTGTTCGCTGGGATTGATACCAAC 3041
 954 PheThrValThrSerLysAlaAsnAspGluGluGlyAspPheAlaGluLeuTyrMetPhe 973
 3042 TTCAGATTACCTCAAGGCATCTGATGAGGATGGGACTTTTGTGAGCTATATGTGTTTC 3101
 974 LysTrpThrThrLeuLeuIleProProThrThrIleLeuIleIleAsnMetValGlyVal 993
 3102 AAGTGGACCAAGTTTGTCTATTCTCGACCACTTAAACAGTGGCTACCAATCCTGGGCTCGCTCTTTGGA 3221
 994 ValAlaGlyThrSerTyrAlaIleAsnSerGlyTyrGlnSerTrpGlyProLeuPheGly 1013
 3162 GTGGCAGGAATTCTTATGCCATTAAACAGTGGCTACCAATCCTGGGCTCGCTCTTTGGA 3221
 1014 LysLeuPhePheAlaPheTrpValIleValHisLeuTyrProPheLeuLysGlyLeuMet 1033
 3222 AAGCTGTTCTTCTCGATCTGGGTGATCTCCATCTCTACCCCTTCTCAAGGGTCTCATG 3281
 1034 GlyArgGlnAsnArgThrProThrIleValIleValIleValIleValIleValIleValIle 1053
 3282 GGAAGGCAGAACCGCACCAACATCGTCAATGTCTGGTCCATCTCTTTCATCTATC 3341
 1054 PheSerLeuLeuTyrValArgValAspProPhe-----ThrThrArgLeuAlaGly 1070
 3342 TTCCTCTGCTGGGTGAGATCGATCTTTCATCTCCCGCACACAGAAAGCTGTGCC 3401
 1071 ProAsnIleGlnThrCysGlyIleAsnCys 1080
 3402 -----TTGGGCAATGTGGCGTCAACTGC 3425

Search completed: August 22, 2004, 17:53:11
 Job time : 1096 secs